Natural variation of molecular and morphological responses to genetic perturbations in Arabidopsis

Youn-Jeong Nam

Natural variation of molecular and morphological responses to genetic perturbations in Arabidopsis

Youn-Jeong Nam

UNIVERSITEIT GENT

VIB

Plant Systems Biology
A VIB-UGENT DEPARTMENT

Ghent University
Faculty of Science
Department of Plant Biotechnology and Bioinformatics
VIB, Plant System Biology

Natural variation of molecular and morphological responses to genetic perturbations in Arabidopsis

Youn-Jeong Nam

Promoter: **Prof. Dr. Dirk Inzé**
Co-promoter: **Dr. Nathalie Gonzalez**

Thesis submitted in partial fulfillment of the requirement for the degree of
Doctor of Philosophy (Ph.D.) in Science: Biochemistry and Biotechnology
Academic year 2015-2016

Examination commission

Chair:

Prof. Dr. Geert De Jaeger

Department of Plant Systems Biology, Faculty of Science, Ghent University
Department of Plant Biotechnology and Bioinformatics, VIB

Secretary:

Prof. Dr. Moritz Nowack

Department of Plant Systems Biology, Faculty of Science, Ghent University
Department of Plant Biotechnology and Bioinformatics, VIB

Members:

Prof. Dr. Dirk Reheul

Department of Plant Production, Faculty of Bioscience Engineering, Ghent University

Prof. Dr. Tom Beeckman
Department of Plant Systems Biology, Faculty of Science, Ghent University
Department of Plant Biotechnology and Bioinformatics, VIB

Dr. Eunyoung Chae
Department of Molecular Biology, Max Planck Institute for Developmental Biology

Prof. Dr. Filip Vandenbussche
Department of Physiology, Faculty of Science, Ghent University

Prof. Dr. Isabel Roldan-ruiz
The Institute for Agricultural and Fisheries Research (ILVO)
Faculty of Science, Ghent University

Dr. Hilde Nelissen
Department of Plant Systems Biology, Faculty of Science, Ghent University
Department of Plant Biotechnology and Bioinformatics, VIB

CONTENTS

List of Abbreviations

35S CaMV promoter	Cauliflower Mosaic Virus 35S promoter
ABA	Abscisic acid
adj-P	adjusted P value
amiPPD	*artificial microRNA PEAPOD*
An-1	Antwerpen
AN3	ANGUSTIFOLIA 3
ANT	AINTEGUMENTA
Blh-1	Bulhary
CDK	Cylin-dependent kinase
cDNA	complementary DNA
CDS	Coding sequence
CK	Cytokinin
Col-0	Columbia
Cot	Cotyledon
Cvi-0	Cape Verde Islands
CYC	Cylin
DAR	DA1-RELATED
DAS	Days after stratification
DE gene	Differentially expressed gene
DNA	Deoxyribonucleic acid
DN-DA1-OE	Overexpression of dominant negative form of *DA1*
EI	Endoreduplication index
Ey15-2	Eyach
GA	Gibberellin
GA20ox-OE	*GIBBERELLIN 20 OXIDASE -OVEREXPRESSION*
GRF	GROWTH-REGULATING FACTOR
GWAS	Genome wide association study
JA	Jasmonic acid

L	Leaf
Ler-0	Landsberg erecta
LS	Leaf series analysis
MIRGIS	Multi-camera *in vivo* rosette growth imaging system
Oy-0	Oystese
p	P value
PRA	Projected rosette area
qRT-PCR	quantitative reverse transcription polymerase chain reaction
QTL	Quantitative trait locus
r	correlation coefficient
RIL	Recombinant inbred line
RNA	Ribonucleic acid
RNA seq	RNA sequencing
Ros	Rosette-expressivity
SA	Salicylic acid
SAM	Shoot apical meristem
SAUR	SMALL AUXIN UPREGULATED RNA
Sel	Selective-expressivity
Sha	Shakdara
T_1	The first generation of transgenic line
T_2	The second generation of transgenic line
T_3	The third generation of transgenic line
TF	Transcription factor
Ts	Transgenics
WalhaesB4	Walddorf-Haslach
wt	wild type
WUS	WUSCHEL
Yeg-1	Yeghegis

Part 1. Introduction

Chapter 1

Leaf growth in Arabidopsis

Youn-Jeong Nam[1,2], Dirk Inzé[1,2*] & Nathalie Gonzalez[1,2*]

[1]Department of Plant Systems Biology, VIB B-9052 Gent, Belgium, [2]Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Gent, Belgium
* These authors contributed equally to this work.

Contributions: Y.J.N. was the main author of this chapter. D.I. and N.G. contributed to the writing of this chapter.

Chapter 1. Leaf growth in Arabidopsis

Leaves are very important organs for plants. Through photosynthesis, they capture light energy and convert it into chemical energy to fuel plants activities throughout their life span. Leaves are also able to sense diverse environmental factors in order to adapt to changing growth conditions, especially those related to light quality and quantity. Furthermore, other plant aboveground parts, such as flowers, bracts, and needles are derived from leaves or modified leaves (Zhao, 2010; Davies, 2012; Rai et al., 2015). Therefore, the study of leaf growth and development can provide a better understanding of shoot development of plants in general. In addition, world's demand for plant-derived products such as food, feed and bio-energy products is gradually increasing. Thus, understanding how leaf growth is regulated represents a crucial research topic in order to generate crops with higher yield.

Arabidopsis is one of the model organisms in plant biology. It is commonly used in laboratory for genetic and molecular biology studies because of several advantages such as a short life span, a small plant size and the existence of a wide range of varieties from around world. Furthermore Arabidopsis contains a small genome that was the first plant genome sequenced (The Arabidopsis Genome Initiative, 2000).

1.1 Arabidopsis leaf development: different tightly regulated phases

Arabidopsis leaf originates from the shoot apical meristem (SAM) where the growth of all aboveground organs initiates. After primordium emergence from the SAM, different developmental phases occur before the leaf reaches its final size and shape: a cell proliferation phase, a transition phase during which cell division and expansion co-occur in the leaf, a cell expansion phase, a meristemoid division phase, and a differentiation phase. Numerous genes were identified that regulate the different stages of leaf development/growth, their mutations leading to an alteration of final leaf size. Some of those genes will be discussed below and in Chapter2.

1.1.1 Shoot Apical Meristem: Stem cell maintenance and leaf initiation

The shoot apical meristem (SAM) is the source of all aerial parts of the plants. In dicots, the SAM consists of three functional zones (central zone (CZ), peripheral zone (PZ), and rib zone (RZ)) (Kwiatkowska, 2008) and three layers (epidermal (layer 1), sub epidermal (layer 2), and inner layer (layer 3)) (Figure 1.1). The CZ, where cells stay at undifferentiated state, is surrounded by PZ. The cells in the PZ have a relatively fast division rate and will differentiate into aboveground plant parts. The RZ is located under the central zone and gives rise to the central tissues of the stem (Figure 1.1).

Stem cells, undifferentiated dividing cells that are capable of differentiating into specialized cells, are present in the center of layer1, layer2 and layer3 and their maintenance is

tightly regulated at genetic level by WUSCHEL (WUS) and CLAVATA (CLV). WUS, an homeodomain transcription factor, is expressed in the RZ and controls stem cell fate by making new stem cells actively divide while old ones lose their stem cell ability (Yadav and Reddy, 2011). *CLV* genes products, CLV 1, 2, and 3, are expressed in the CZ and act as receptors (CLV1 and CLV2) and ligand (CLV3) (Brand et al., 2000; Schoof et al., 2000). In the WUS/CLV signalling pathway, WUS and CLV3 act as mobile signals. WUS, expressed in layer 3 cells, activates the expression of *CLV3* in the stem cells in a non-cell autonomous manner. CLV3 inhibits WUS activity after binding to the receptors, CLV1 or CLV2, in layer 3 (Haecker and Laux, 2001; Braybrook and Kuhlemeier, 2010). This WUS/CLV loop regulation is suppressed by the F-box protein LEAF CURLING RESPONSIVENESS (LCR) that inhibits WUS signal. LCR does not target the activity or stability of WUS, but probably another protein which is involved in modifying or assisting WUS for its function or movement (Knauer et al., 2013). miR394, a mobile micro RNA produced in layer1, positively regulates WUS via repressing *LCR* (Knauer et al., 2013) (Figure 1.1).

It has been shown that the phytohormone cytokinin (CK) is involved in SAM maintenance by inducing the expression of *WUS* (Leibfried et al., 2005; Gordon et al., 2009) which in turn represses *ARABIDOPSIS type-A RESPONSE REGULATORS* (*ARRs*), negative regulators of cytokinin signalling (Figure 1.1). Interestingly, Arabidopsis plants over-expressing cytokinin dehydrogenase genes (Werner et al., 2003) or mutated in cytokinin receptors (Higuchi et al., 2004) produce smaller meristems but also smaller leaves suggesting a relation between meristem size and leaf size. Stem cell capacity is maintained by a strict balance between the two phytohormones cytokinin and gibberellin (GA). High levels of CK and low levels of GA suppress cell differentiation in the SAM (Gordon et al., 2009; Veit, 2009). KNOTTED LIKE HOMEOBOX (KNOX1), a homeodomain transcription factor, maintains stem cells by regulating the antagonistic action of these two phytohormones. CK and GA levels are positively and negatively regulated by KNOX1, respectively, through the modulation of their biosynthesis enzymes: ISO-PENTENYLTRANSFEREASE7 for CK and GA20ox for GA (Sakamoto et al., 2001; Jasinski et al., 2005). ASYMMETRIC LEAF1/ROUGH SHEATH2/PHANTASTICA (ARP) are transcription factors having an antagonistic role with KNOX. *ASYMMETRIC LEAVES 1* (*AS1*) is a member of the *ARP* family expressed in primordia forming cells restricting the expression of *KNOX* in the meristematic cells (Figure 1.2).

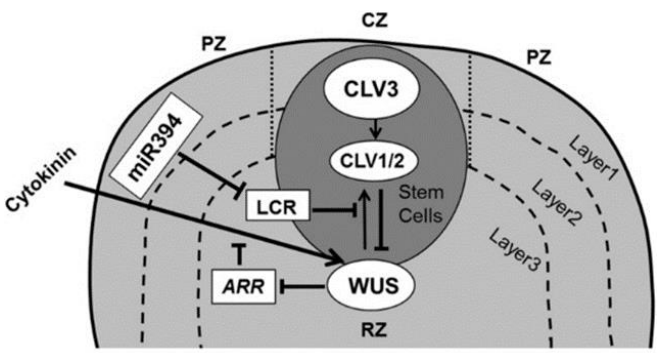


Figure 1.1 Structure of the SAM and maintenance of stem cells in the SAM. The SAM consists of three functional zones (central zone (CZ), peripheral zone (PZ), and rib zone (RZ)) and three layers. In the three layers, an antagonistic regulatory loop between WUS and CLV preserves stem cells in the meristem. Other regulators involved in the regulation of WUS/CLV are Cytokinin, ARR, LCR, and *miR394* (from Kalve et al., 2014).

High accumulation of auxin at the periphery of the SAM induces development of leaf primordia and suppresses KNOX1 activity. The activity of the auxin influx carrier, AUXIN RESISTANT (AUX), and the efflux transporter proteins, PIN-FORMED1 (PIN1) contributes to this accumulation of auxin. Modeling studies suggest that AUX and PIN1 create an auxin movement flow along the layer 1 in the primordia and then a flow towards the base of the shoot from the tip of the primordium to stimulate venation differentiation in layer 2 and 3 (Reinhardt et al., 2003; de Reuille et al., 2006) (Figure 1.2).

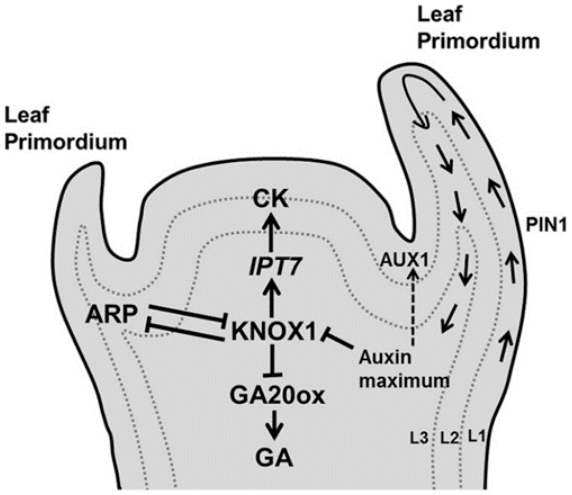


Figure 1.2 Mechanism of stem cell maintenance and leaf initiation. KNOX1 plays an important role for stem cell maintenance by balancing the levels of the two phytohormones, GA and CK, in the SAM (from Kalve et al., 2014).

The mechanism of how cells differentiating into a leaf are distinguished from the rest of the stem cells has been described. Two specific transcription factors families are involved in cell identity maintenance: AS1 and CUP-SHAPED COTYLEDON (CUC) 1, 2, and 3 that are members of the MYB and NAC families, respectively. In the leaf primordium, AS1 forms a complex with AS2 to repress the expression of *KNOX1* genes (Guo et al., 2008). CUC proteins are expressed in specific cells between the meristem and the leaf primordium and CUC1 positively regulates the *KNOX1* expression (Aida et al., 1999; Takada et al., 2001; Vroemen et al., 2003). miR164 regulates the expression of *CUC* to limit their expression in specific tissues (Laufs et al., 2004; Mallory et al., 2004).

In summary, KNOX1 that is regulated by ARP and auxin is necessary to maintain stem cells in the SAM by balancing the accumulation of two phytohormones, CK and GA, and its down-regulation by AS1-AS2 allows the initiation of the leaf primordium.

1.1.2 Cell proliferation phase

After emergence from the SAM, the leaf primordium grows mainly through cell proliferation leading to an increase in cell number. Cell proliferation occurs first in the entire leaf primordium and after few days, cells stop dividing from the tip to the bottom of the leaf forming a cell division gradient. The cell cycle, corresponding to the series of events leading to DNA duplication (DNA replication or synthesis) and cell division through mitosis to produce two

daughter cells, is divided in four phases, Gap 1 (G1), Synthesis (S), Gap 2 (G2) and Mitosis (M) phases. The cell cycle is strictly controlled by molecular components which are highly conserved in eukaryotes (Inzé and De Veylder, 2006; Gutierrez, 2009). More specifically, the cell cycle is regulated by the activity of CDK/CYC complexes consisting of a CYCLIN-DEPENDENT KINASE (CDK) as catalytic subunit and a CYCLIN (CYC) as regulatory subunit. During the cell cycle these complexes get activated via phosphorylation by CDK ACTIVATING KINASES (CAK) while inhibited by KIP RELATED PROTEINS (KRPs) also known as INHIBITOR of CDK (CKI). CDKA/CYCD complexes consisting out of D-TYPE CYCLIN (CYCD) and A-TYPE CYCLIN dependent kinase (CDKA) are central for the G1/S phase transition, activating DNA duplication. After DNA duplication, the cell enters the G2 phase to prepare for division through mitosis (G2/M phase transition). CDKA and CDKB as well as CYCA, CYCB, and CYCD are involved in these processes. The exit from mitosis requires ubiquitin-dependent proteolytic degradation of CYCs which is mediated by the ANAPHASE-PROMOTING COMPLEX/CYCLOSOME (APC/C), a large multi protein E3 ubiquitin ligase (Sullivan and Morgan, 2007; Marrocco et al., 2010). Overexpression of *APC10*, encoding a subunit of the APC/C complex, leads to the production of larger leaves with more cells as a result of increased cell cycle rate (Eloy et al., 2011). On the other hand, a knock out mutant of *SAMBA*, a subunit of APC/C and negative regulator of the cell cycle, shows larger meristem, seeds, leaves, and roots (Eloy et al., 2012). In contrast, plants overexpressing *KRP1* present reduced leaf size with decreased number of cells (Weinl et al., 2005). It has been shown that DELLAs, repressors of GA signalling, restrain cell proliferation by promoting the expression level of two *CKI*, *KPR2* and *SIAMESE* (*SIM)* and that GA signalling positively regulates cell proliferation by inducing the degradation of DELLAs (Achard et al., 2009).

1.1.3 Transition from cell division to cell expansion

During leaf development, a transition phase from cell division to cell expansion happens during which a cell cycle arrest front moves rapidly, in a not so gradual manner, through the leaf from the tip to the base (Andriankaja et al., 2012) (Figure 1.3A). A number of regulators have been found to control the duration of the cell proliferation phase and therefore to affect final leaf size showing that this transition is important for the regulation of leaf development.

Auxin induces the expression of AUXIN-REGULATED GENE INVOLVED IN ORGAN SIZE (ARGOS) that regulates this transition phase (Figure 1.3B). Transgenic plants overexpressing *ARGOS* produce enlarged aerial organs whereas plants expressing an antisense sequence of *ARGOS* display reduced organ size. The alteration of organ size results from changes in cell number and the duration of growth (Hu et al., 2003). ARGOS might influences the expression of *AINTEGUMENTA* (*ANT*), a DNA binding protein, and *CYCD3;1,* a G1-type cyclin, proteins controlling cell cycle initiation and progression, respectively (Hu et al., 2003). The increased leaf size in *ARGOS* overexpressing plant can be inhibited by loss of function of *ANT* (Hu et al., 2003). Ectopic

expression of *ANT* and *CYCD3;1* leads to production of large leaves due to an increased cell number resulting from a prolonged duration of cell proliferation (Hu et al., 2003). A recent work has revealed that ARGOS also functions as a negative feedback regulator of ethylene signalling (Rai et al., 2015). *ARGOS* is one member of the *ARGOS* gene family and the expression of the members of this family is also induced by ethylene. The induction in expression requires signaling through the primary ethylene signaling pathway, however, it is repressed in ethylene-insensitive mutants. Seedlings overexpressing *ARGOS* family members, *ARGOS* and A*RGOS-LIKE (ARL), ARL2,* and *ARL3*, show reduced ethylene sensitivity and have similar phenotype with ethylene-insensitive mutants. It was shown that *ARGOS* family regulates ethylene-dependent gene expression (Rai et al., 2015).

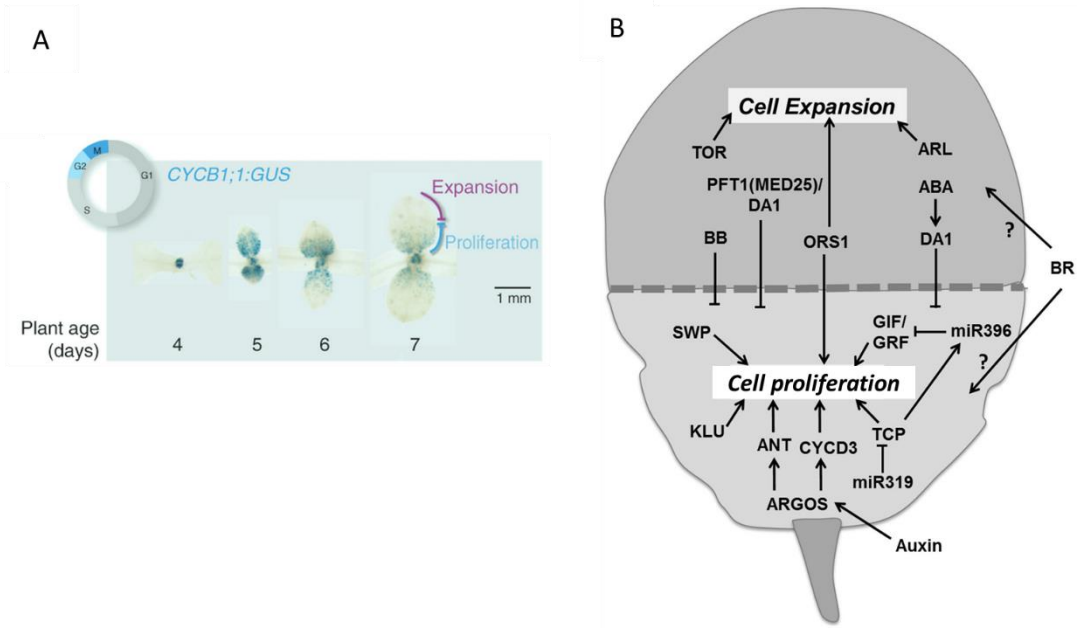


Figure 1.3 Regulation of cell proliferation and cell expansion during leaf development. (A) Localization of the cell proliferation region in developing leaves as plants get older. A CYCLINB1;1, (CYCB1;1:GUS) reporter marks the cells in the G2-M phase of the cell cycle (from Rodriguez et al., 2014). (B) Molecular mechanisms regulating the transition between cell proliferation and cell expansion. The transition part is highlighted with the dashed line and the proliferation and cell expansion zones are below and above this line, respectively (from Kalve et al., 2014).

Transcription factors such as GROWTH REGULATING FACTOR (GRF) and TEOSINTE BRANCHED1/CYCLOIDEA/PCF4 (TCP4) are also key regulators of the transition between cell proliferation and cell expansion (Figure 1.3B). GRFs belong to a plant-specific transcription factor family that contains nine members in Arabidopsis and that regulates cell proliferation or cell expansion (Kim et al., 2003; Horiguchi et al., 2005). For example, *GRF5* triggers a larger leaf size phenotype when overexpressed while loss of function plants produce smaller organs (Lee et al., 2009). GRF5 interacts with a growth regulator, GRF-INTERACTING FACTOR1 (GIF1)/ANGUSTIFOLIA3 (AN3)(Horiguchi et al., 2005) and the expression of *GRF* and *GIF* is

negatively controlled by miR396 (Liu et al., 2009). TCP4, a transcription factor member of the TCP family, regulates the expression of *miR396* (Rodriguez et al., 2010) and is negatively regulated by *miR319* (Palatnik et al., 2003). In a mutant of *miR319*, TCP4 activates *miR396* and this results in a decrease leaf size through the inhibition of *GRF* expression. Similarly, overexpressing *TCP4* causes reduced leaf size (Rodriguez et al., 2010). Not only *TCP4* but also other members of the *TCP* family are targets of *miR319* and regulate cell division arrest front in Arabidopsis (Palatnik et al., 2003; Efroni et al., 2008). Plants overexpressing *miR319* show highly reduced expression of *TCP2*, *TCP3*, *TCP4*, *TCP10*, and *TCP24* and produce larger and crinkle leaves (Palatnik et al., 2003; Schommer et al., 2008).

In addition to regulation of transcription, other leaf growth regulators belonging to different functional categories are important to promote the transition phase between division and expansion. For example, *KLUH (KLU)/CYP78A5* encoding a cytochrome P450 monooxygenase produces a yet unidentified mobile growth factor promoting cell proliferation (Anastasiou et al., 2007) (Figure 1.3B). Overexpressing *KLU* from its endogenous promoter causes the formation of larger organs as a result of increased cell proliferation whereas mutation of *KLU* produces smaller leaves (Anastasiou et al., 2007; Kazama et al., 2010; Stransfeld et al., 2010).

1.1.4 Cell expansion and endoreduplication

During the transition phase and after the end of cell proliferation at the base of the leaf, cells start expanding in the whole leaf. Cell expansion is driven by turgidity, the pressure caused by osmotic flow of water, and its direction and extent are influenced by the biomechanical properties of cell walls (Bashline et al., 2014). EXPANSIN (EXP) proteins are located in the plant cell wall to allow its loosening (Cosgrove, 2005). Plants overexpressing EXP10 display larger leaf blades with larger cells whereas down-regulation of its expression causes the formation of smaller rosette (Cho and Cosgrove, 2000).

Hormone signalling also affects cell expansion. For example, the auxin-induced SMALL AUXIN UP-RNA (SAUR) proteins have a role as positive regulators of cell expansion (Spartz et al., 2012; Kong et al., 2013; Stamm and Kumar, 2013; Spartz et al., 2014). Auxin rapidly activates plasma membrane H^+-ATPases to lower pH, in turn activating expansins and other cell wall–modifying proteins such as XYLOGLUCAN ENDOTRANSGLYCOSIDASEs (XET) to mediate cell expansion (Takahashi et al., 2012). This activation of plasma membrane H^+-ATPases by auxin is mediated by SAUR19. SAUR19 physically interacts with and inhibits PP2C-D phosphatases, a family of type 2C protein phosphatases that are negative regulators of plasma membrane H^+-ATPases. Thus, SAUR19 promotes cell expansion (Spartz et al., 2014). ARGOS-LIKE (ARL) which shows sequence homology with ARGOS is also involved in leaf growth regulation (Hu et al., 2006). ARL is induced by brassinosteroid and regulates cell expansion. Overexpression of *ARL* produces larger leaves while its loss of function displays smaller leaf phenotype (Hu et al., 2006).

When cell division stops, some cells continue replicating their nuclear genome in the absence of mitosis. This phenomenon is termed endoreduplication which is a common process in plants whereas it is observed only in specific cell types in animals (Galbraith et al., 1991). Endoreduplication causes high ploidy levels in the cell and it has been reported that these high ploidy levels positively correlate with cell size in plants (Sugimoto-Shirasu and Roberts, 2003; Breuer et al., 2007; Tsukaya, 2013). To begin endoreduplication, CDK activity has to be kept low enough (De Veylder et al., 2011). In transgenic plants overexpressing a dominant negative *CDKB1;1* cell cycle progression is inhibited and blocked in G2 phase (Boudolf et al., 2004). CDKB1;1 forms an active complex when interacting with CYCA2;3 that suppresses endoreduplication in the leaf (Boudolf et al., 2009). Loss of function mutants of *CYCA2;3* produce leaves having increased ploidy levels (Imai et al., 2006).

In plants, the CCS52A proteins function in the regulation of the switch from mitotic to endoreduplication cycles controlling the number of mitotic cells and the endoreduplication events. It has been shown that elevated ectopic CCS52A expression positively correlates with organ size (Baloban et al., 2013). Also significant positive correlation was found between high endopolyploidy and large leaf size through a survey of endopolyploidy, cell number, and cell size in leaves on *Arabidopsis* accessions (Gegas et al., 2014).

1.1.5 Compensation: interconnection between cell division and cell expansion

During leaf development, cell proliferation and cell expansion are tightly interconnected. It has been observed in several mutants that a reduced cell number is compensated with enlarged cells (Hisanaga et al., 2015). For instance, the decreased cell number in *gif* loss of function or *KIP-RELATED PROTEIN2* (*KRP2*) gain of function is associated with an increased cell size that partially compensates the reduction in organ size (Hemerly et al., 1995; De Veylder et al., 2001; Horiguchi et al., 2005; Ferjani et al., 2007). Therefore, this compensation phenomenon highlights the coordinated regulation of the cell division and cell expansion phase (Hisanaga et al., 2015).

Compensating cell enlargement can be divided in 3 modes of action based on the kinetic of cell size increase and three classes of mutants have been described (Ferjani et al., 2007). Class 1 mutants such as *gif* display an increased rate of cell expansion, while class 2 mutants have an increased duration of cell expansion. *fugu5* mutant, corresponding to the loss of function of *FUGU5* encoding a H(+)-translocating inorganic pyrophosphatase is an example of class 2 mutant. The mutants belonging to class 3, such as plants overexpressing *KRP2,* show proliferating cells that are already larger than the wild type, and also show an increased cell expansion rate.

1.1.6 Meristemoid division

In Arabidopsis, final leaf size is controlled by complex molecular mechanisms which govern interlinked growth phases including a phase of meristemoid division (Gonzalez et al., 2012) (Figure 1.4A). Meristemoids are stem cell-like cells, dispersed in the leaf epidermis, that divide asymmetrically to give rise to a pavement cell and a meristemoid cell.

Guard cell development is initiated by an asymmetric cell division of a protodermal cell (Figure 1.4B). The two daughter cells obtain different identities; one maintains protodermal cell identity, whereas another one becomes a meristemoid mother cell (MMC). The MMC divides asymmetrically to produce a larger stomatal lineage ground cell (SLGC) and smaller meristemoid. A SLGC can terminally differentiate into a lobed pavement cell or alternatively, a SLGC can initiate an asymmetric spacing division to produce a satellite meristemoid, which is always placed distal to an existing stoma or precursor. A meristemoid can undergo, at a maximum, three sequential divisions, called amplifying division, therefore generating three pavement cells before differentiating into a guard mother cell (GMC). GMC divides once symmetrically, producing paired guard cells (GCs) that will form a stoma. 67 % and 48 % of all pavement cells originate from the division of meristemoids in cotyledons and leaves, respectively (Geisler et al., 2000; Bergmann and Sack, 2007). Meristemoid division therefore contributes to producing a large proportion of pavement cells.

Stomatal fate is determined by three helix-loop-helix (bHLH) transcription factors, SPEECHLESS (SPCH), MUTE, and FAMA. SPCH regulates conversion of a protodermal cell into a MMC (MacAlister et al., 2007). MUTE controls the transition from meristemoid to GMC (Pillitteri et al., 2007) and FAMA is essential to make functional guard cells from GMC (Ohashi-Ito and Bergmann). FAMA needs to bind to another protein called RETINOBLASTOMA-RELATED (RBR) to function properly. It has been suggested that these two proteins make the transition from stem cell to guard cell permanent by changing the structure of DNA in regions that control stem cell genes (Matos et al., 2014).

PEAPOD (*PPD*) genes encoding transcriptional regulators are known to control meristemoid division (White, 2006). The mode of action of PPD will be discussed in detail in the chapter 2.

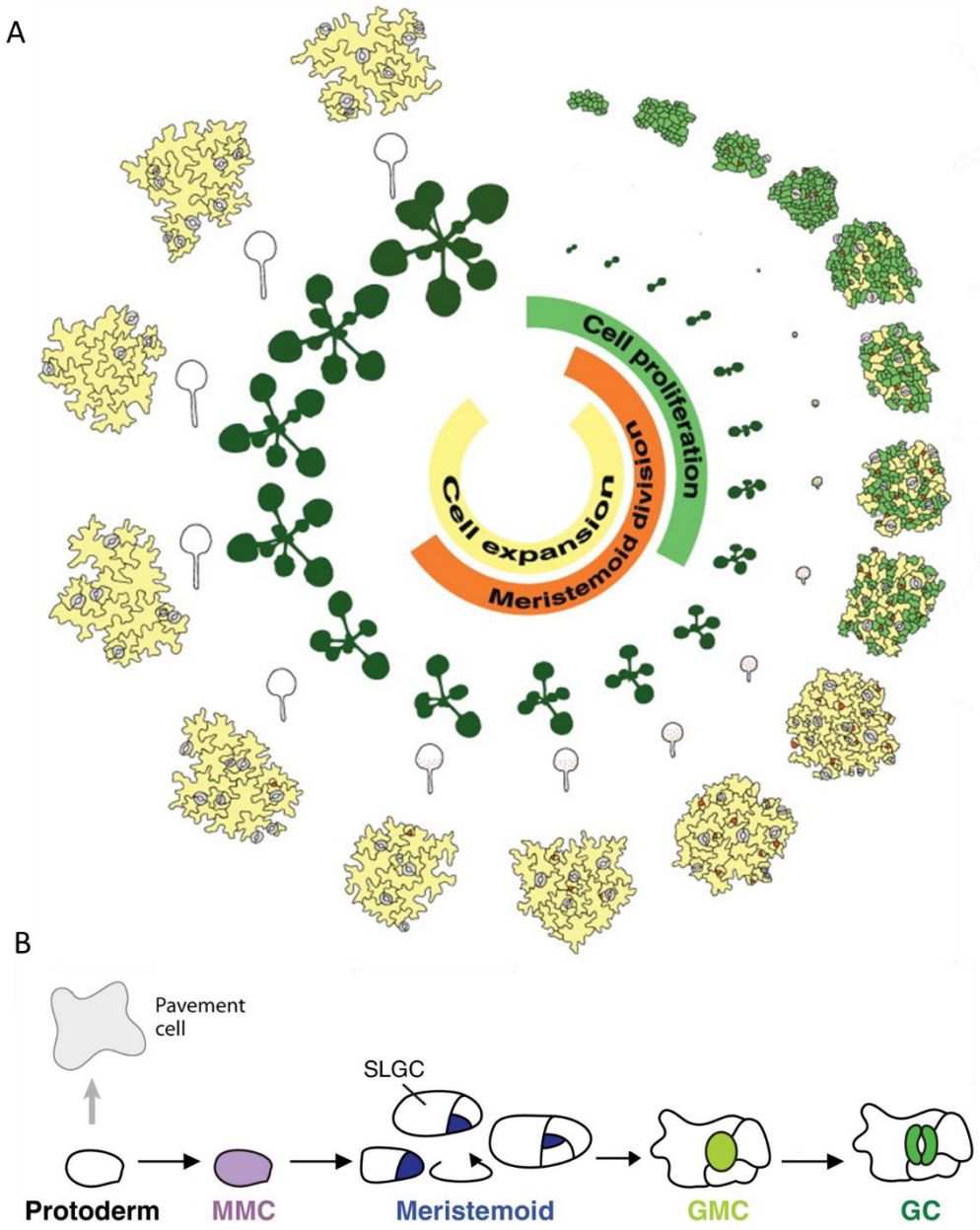


Figure 1.4 Different developmental stages at rosette, leaf and cellular level and stomatal development in Arabidopsis. (A) Different phases of development (cell proliferation phase, meristemoid division phase and cell expansion phase). The images of the rosette, of the first leaf pair, and of the cell drawings are from 4 to 20 days old plants. Dividing cells in the primary general cell-division front are represented in green, meristemoid cells in orange and expanding cells in yellow. Cell drawings were made from the abaxial side of the leaf epidermis (from Gonzalez et al., 2012). (B) Stomatal development. A protodermal cell (white) will become a meristemoid mother cell (MMC) (purple) while other protodermal cells differentiate into pavement cells. To enter the stomatal lineage, MMC undergo a first asymmetric division producing a meristemoid (dark blue) and a stomatal lineage ground cell (SLGC). After sequential asymmetric amplifying divisions, meristemoid differentiate into a guard mother cell (GMC) (light green). The GMC divides once symmetrically to form two guard cells (GCs) (green) that will differentiate into a stoma (modified from Torii, 2015).

1.1.7 Differentiation

During leaf development, all cell types are generated from undifferentiated stem cells and then differentiate to achieve specific morphological, biochemical and physiological functions. For instance, the leaf is composed of mesophyll tissue, vascular tissue, pavement cells, guard cells, stomata and trichomes.

High local concentration of auxin at the periphery of the SAM induces not only leaf initiation but also provascular identity during early leaf developmental stage. This auxin dynamic leads to differentiation of procambial cells which are vascular stem cells into the midvein and lateral veins. This process requires enhanced expression of ATHB8 (ARABIDOPSIS HOMEOBOX TRANSCRIPTION FACTOR 8), a positive regulator of venation differentiation (Scarpella et al., 2004; Scarpella et al., 2006; Bayer et al., 2009). A correlation between leaf size and venation patterns has been observed across 485 globally distributed species (Sack et al., 2012) suggesting a link between these two events.

1.1.8 Heteroblasty

In this thesis, for quantification of growth of the aboveground part of the plant, all rosette leaves of Arabidopsis were taken into account. Therefore, it is important to know that a plant generates different individual leaves which do not have similar shape and size.

Heteroblasty corresponds to the effects of developmental phases of the plant particularly in relation to the size and morphology of the leaves (Zotz et al., 2011) (Figure 1.5). During vegetative development, Arabidopsis plants go first through a juvenile phase, vegetative growth following germination, during which newly produced leaves have trichomes only on their adaxial side. The juvenile phase is followed by an adult phase after which the plant is capable to produce flowers. During the adult phase, the newly produced leaves start to have trichomes also on their abaxial side, they get more elongated and their margins have deeper serrations than juvenile leaves (Figure 1.5). The molecular mechanism regulating the transition between a juvenile and an adult leaf involves the miR156 related network (Wang et al., 2008; Wang et al., 2009). Expression of miR156 is necessary and sufficient to maintain the juvenile phase by inhibiting the expression of SQUAMOSA PROMOTER BINDING PROTEINLIKE (SPL) transcription factors which promote the adult traits (Schmid et al., 2003; Wang et al., 2008; Wang et al., 2009). As the plant ages, the expression *miR156* decreases leading to an increase in the expression of *SPL*.

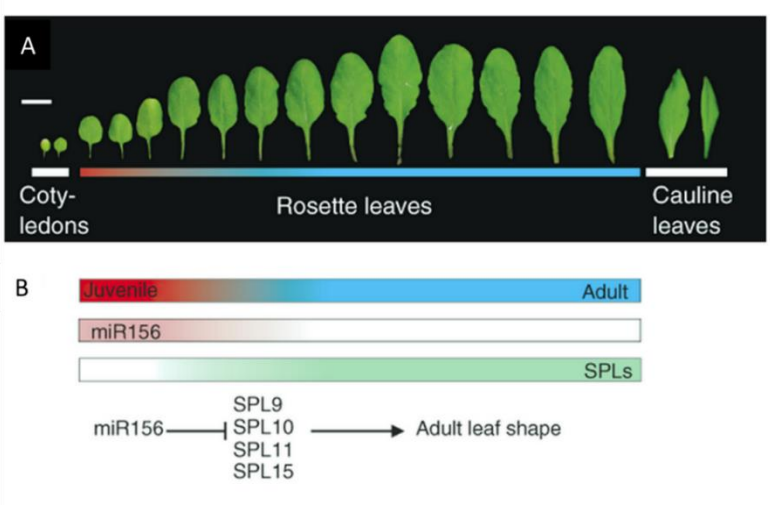


Figure 1.5 Heteroblasty in Arabidopsis. (A) Leaf series of Arabidopsis showing cotyledons, successive true rosette leaves and cauline leaves ordered by developmental ages. (B) Molecular regulation of heteroblasty in rosette leaves during vegetative phase by antagonistic action of *miR165* and *SPL* (from Rodriguez et al., 2014).

References

Achard, P., Gusti, A., Cheminant, S., Alioua, M., Dhondt, S., Coppens, F., Beemster, G.T.S., and Genschik, P. (2009). Gibberellin signaling controls cell proliferation rate in *Arabidopsis*. Curr. Biol. **19:** 1188-1193.

Aida, M., Ishida, T., and Tasaka, M. (1999). Shoot apical meristem and cotyledon formation during Arabidopsis embryogenesis: interaction among the CUP-SHAPED COTYLEDON and SHOOT MERISTEMLESS genes. Development **126:** 1563-1570.

Anastasiou, E., Kenz, S., Gerstung, M., MacLean, D., Timmer, J., Fleck, C., and Lenhard, M. (2007). Control of plant organ size by *KLUH/CYP78A5*-dependent intercellular signaling. Dev. Cell **13:** 843-856.

Andriankaja, M., Dhondt, S., De Bodt, S., Vanhaeren, H., Coppens, F., De Milde, L., Mühlenbock, P., Skirycz, A., Gonzalez, N., Beemster, G.T.S., and Inzé, D. (2012). Exit from proliferation during leaf development in *Arabidopsis thaliana*: a not-so-gradual process. Dev. Cell **22:** 64-78.

Baloban, M., Vanstraelen, M., Tarayre, S., Reuzeau, C., Cultrone, A., Mergaert, P., and Kondorosi, E. (2013). Complementary and dose-dependent action of AtCCS52A isoforms in endoreduplication and plant size control. New Phytol. **198:** 1049-1059.

Bashline, L., Lei, L., Li, S., and Gu, Y. (2014). Cell wall, cytoskeleton, and cell expansion in higher plants. Mol. Plant **7:** 586-600.

Bayer, E.M., Smith, R.S., Mandel, T., Nakayama, N., Sauer, M., Prusinkiewicz, P., and Kuhlemeier, C. (2009). Integration of transport-based models for phyllotaxis and midvein formation. Genes Dev. **23:** 373-384.

Bergmann, D.C., and Sack, F.D. (2007). Stomatal development. Annu. Rev. Plant Biol. **58:** 163-181.

Boudolf, V., Barrôco, R., de Almeida Engler, J., Verkest, A., Beeckman, T., Naudts, M., Inzé, D., and De Veylder, L. (2004). B1-type cyclin-dependent kinases are essential for the formation of stomatal complexes in *Arabidopsis thaliana*. Plant Cell **16:** 945-955.

Boudolf, V., Lammens, T., Boruc, J., Van Leene, J., Van Den Daele, H., Maes, S., Van Isterdael, G., Russinova, E., Kondorosi, E., and Witters, E. (2009). CDKB1; 1 forms a functional complex with CYCA2; 3 to suppress endocycle onset. Plant Physiol. **150:** 1482-1493.

Brand, U., Fletcher, J.C., Hobe, M., Meyerowitz, E.M., and Simon, R. (2000). Dependence of stem cell fate in *Arabidopsis* on a feedback loop regulated by *CLV3* activity. Science **289:** 617-619.

Braybrook, S.A., and Kuhlemeier, C. (2010). How a plant builds leaves. Plant Cell **22:** 1006-1018.

Breuer, C., Stacey, N.J., West, C.E., Zhao, Y., Chory, J., Tsukaya, H., Azumi, Y., Maxwell, A., Roberts, K., and Sugimoto-Shirasu, K. (2007). BIN4, a novel component of the plant DNA topoisomerase VI complex, is required for endoreduplication in Arabidopsis. The Plant Cell **19:** 3655-3668.

Cho, H.-T., and Cosgrove, D.J. (2000). Altered expression of expansin modulates leaf growth and pedicel abscission in Arabidopsis thaliana. Proceedings of the National Academy of Sciences **97:** 9783-9788.

Cosgrove, D.J. (2005). Growth of the plant cell wall. Nat. Rev. Mol. Cell Biol. **6:** 850-861.

Davies, P. (2012). Plant hormones and their role in plant growth and development. (Springer Science & Business Media).

de Reuille, P.B., Bohn-Courseau, I., Ljung, K., Morin, H., Carraro, N., Godin, C., and Traas, J. (2006). Computer simulations reveal properties of the cell-cell signaling network at the shoot apex in Arabidopsis. Proc Natl Acad Sci U S A **103:** 1627-1632.

De Veylder, L., Larkin, J.C., and Schnittger, A. (2011). Molecular control and function of endoreplication in development and physiology. Trends Plant Sci. **16:** 624-634.

De Veylder, L., Beeckman, T., Beemster, G.T., Krols, L., Terras, F., Landrieu, I., Van Der Schueren, E., Maes, S., Naudts, M., and Inzé, D. (2001). Functional analysis of cyclin-dependent kinase inhibitors of Arabidopsis. The Plant Cell **13:** 1653-1668.

Efroni, I., Blum, E., Goldshmidt, A., and Eshed, Y. (2008). A protracted and dynamic maturation schedule underlies *Arabidopsis* leaf development. Plant Cell **20:** 2293-2306.

Eloy, N.B., de Freitas Lima, M., Van Damme, D., Vanhaeren, H., Gonzalez, N., De Milde, L., Hemerly, A.S., Beemster, G.T.S., Inzé, D., and Ferreira, P.C.G. (2011). The APC/C *subunit 10* plays an essential role in cell proliferation during leaf development. Plant J. **68:** 351-363.

Eloy, N.B., et al. (2012). SAMBA, a plant-specific anaphase-promoting complex/cyclosome regulator is involved in early development and A-type cyclin stabilization. Proc. Natl. Acad. Sci. USA **109:** 13853-13858.

Ferjani, A., Horiguchi, G., Yano, S., and Tsukaya, H. (2007). Analysis of leaf development in *fugu* mutants of Arabidopsis reveals three compensation modes that modulate cell expansion in determinate organs. Plant Physiol. **144:** 988-999.

Galbraith, D.W., Harkins, K.R., and Knapp, S. (1991). Systemic endopolyploidy in *Arabidopsis thaliana*. Plant Physiol. **96:** 985-989.

Gegas, V.C., Wargent, J.J., Pesquet, E., Granqvist, E., Paul, N.D., and Doonan, J.H. (2014). Endopolyploidy as a potential alternative adaptive strategy for *Arabidopsis* leaf size variation in response to UV-B. J. Exp. Bot. **65:** 2757-2766.

Geisler, M., Nadeau, J., and Sack, F.D. (2000). Oriented asymmetric divisions that generate the stomatal spacing pattern in Arabidopsis are disrupted by the *too many mouths* mutation. Plant Cell **12:** 2075-2086.

Gonzalez, N., Vanhaeren, H., and Inzé, D. (2012). Leaf size control: complex coordination of cell division and expansion. Trends Plant Sci. **17:** 332-340.

Gordon, S.P., Chickarmane, V.S., Ohno, C., and Meyerowitz, E.M. (2009). Multiple feedback loops through cytokinin signaling control stem cell number within the *Arabidopsis* shoot meristem. Proc. Natl. Acad. Sci. USA **106:** 16529-16534.

Guo, M., Thomas, J., Collins, G., and Timmermans, M.C. (2008). Direct repression of KNOX loci by the ASYMMETRIC LEAVES1 complex of Arabidopsis. Plant Cell **20:** 48-58.

Gutierrez, C. (2009). The Arabidopsis cell division cycle. Arabidopsis Book **7:** e0120.

Haecker, A., and Laux, T. (2001). Cell-cell signaling in the shoot meristem. Curr Opin Plant Biol **4:** 441-446.

Hemerly, A., de A Engler, J., Bergounioux, C., Van Montagu, M., Engler, G., Inzé, D., and Ferreira, P. (1995). Dominant negative mutants of the Cdc2 kinase uncouple cell division from iterative plant development. The EMBO journal **14:** 3925.

Higuchi, M., et al. (2004). In planta functions of the Arabidopsis cytokinin receptor family. Proc Natl Acad Sci U S A **101:** 8821-8826.

Hisanaga, T., Kawade, K., and Tsukaya, H. (2015). Compensation: a key to clarifying the organ-level regulation of lateral organ size in plants. J. Exp. Bot. **66:** 1055-1063.

Horiguchi, G., Kim, G.T., and Tsukaya, H. (2005). The transcription factor AtGRF5 and the transcription coactivator AN3 regulate cell proliferation in leaf primordia of Arabidopsis thaliana. The Plant Journal **43:** 68-78.

Hu, Y., Xie, Q., and Chua, N.-H. (2003). The Arabidopsis auxin-inducible gene ARGOS controls lateral organ size. The Plant Cell **15:** 1951-1961.

Hu, Y., Poh, H.M., and Chua, N.-H. (2006). The Arabidopsis *ARGOS-LIKE* gene regulates cell expansion during organ growth. Plant J. **47:** 1-9.

Imai, K.K., Ohashi, Y., Tsuge, T., Yoshizumi, T., Matsui, M., Oka, A., and Aoyama, T. (2006). The A-type cyclin CYCA2;3 is a key regulator of ploidy levels in *Arabidopsis* endoreduplication. Plant Cell **18:** 382-396.

Inzé, D., and De Veylder, L. (2006). Cell cycle regulation in plant development. Annu. Rev. Genet. **40:** 77-105.

Jasinski, S., Piazza, P., Craft, J., Hay, A., Woolley, L., Rieu, I., Phillips, A., Hedden, P., and Tsiantis, M. (2005). KNOX action in Arabidopsis is mediated by coordinate regulation of cytokinin and gibberellin activities. Current biology : CB **15:** 1560-1565.

Kalve, S., De Vos, D., and Beemster, G.T.S. (2014). Leaf development: a cellular perspective. Front. Plant Sci. **5:** 362.

Kazama, T., Ichihashi, Y., Murata, S., and Tsukaya, H. (2010). The mechanism of cell cycle arrest front progression explained by a *KLUH/CYP78A5*-dependent mobile growth factor in developing leaves of *Arabidopsis thaliana*. Plant and Cell Physiology **51:** 1046-1054.

Kim, J.H., Choi, D., and Kende, H. (2003). The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in *Arabidopsis*. Plant J. **36:** 94-104.

Knauer, S., Holt, A.L., Rubio-Somoza, I., Tucker, E.J., Hinze, A., Pisch, M., Javelle, M., Timmermans, M.C., Tucker, M.R., and Laux, T. (2013). A protodermal miR394 signal defines a region of stem cell competence in the Arabidopsis shoot meristem. Dev Cell **24:** 125-132.

Kong, Y., Zhu, Y., Gao, C., She, W., Lin, W., Chen, Y., Han, N., Bian, H., Zhu, M., and Wang, J. (2013). Tissue-specific expression of SMALL AUXIN UP RNA41 differentially regulates cell expansion and root meristem patterning in Arabidopsis. Plant Cell Physiol. **54:** 609-621.

Kwiatkowska, D. (2008). Flowering and apical meristem growth dynamics. J Exp Bot **59:** 187-201.

Laufs, P., Peaucelle, A., Morin, H., and Traas, J. (2004). MicroRNA regulation of the CUC genes is required for boundary size control in *Arabidopsis* meristems. Development **131:** 4311-4322.

Lee, B.H., Ko, J.-H., Lee, S., Lee, Y., Pak, J.-H., and Kim, J.H. (2009). The Arabidopsis GRF-INTERACTING FACTOR gene family performs an overlapping function in determining organ size as well as multiple developmental properties. Plant Physiol. **151:** 655-668.

Leibfried, A., To, J.P., Busch, W., Stehling, S., Kehle, A., Demar, M., Kieber, J.J., and Lohmann, J.U. (2005). WUSCHEL controls meristem function by direct regulation of cytokinin-inducible response regulators. Nature **438:** 1172-1175.

Liu, D., Song, Y., Chen, Z., and Yu, D. (2009). Ectopic expression of miR396 suppresses GRF target gene expression and alters leaf growth in Arabidopsis. Physiol. Plant. **136:** 223-236.

MacAlister, C.A., Ohashi-Ito, K., and Bergmann, D.C. (2007). Transcription factor control of asymmetric cell divisions that establish the stomatal lineage. Nature **445:** 537-540.

Mallory, A.C., Dugas, D.V., Bartel, D.P., and Bartel, B. (2004). MicroRNA regulation of NAC-domain targets is required for proper formation and separation of adjacent embryonic, vegetative, and floral organs. Current biology : CB **14:** 1035-1046.

Marrocco, K., Bergdoll, M., Achard, P., Criqui, M.C., and Genschik, P. (2010). Selective proteolysis sets the tempo of the cell cycle. Curr Opin Plant Biol **13:** 631-639.

Matos, J.L., Lau, O.S., Hachez, C., Cruz-Ramírez, A., Scheres, B., and Bergmann, D.C. (2014). Irreversible fate commitment in the Arabidopsis stomatal lineage requires a FAMA and RETINOBLASTOMA-RELATED module. eLife **3:** e03271.

Ohashi-Ito, K., and Bergmann, D.C. Arabidopsis FAMA Controls the Final Proliferation.

Palatnik, J.F., Allen, E., Wu, X., Schommer, C., Schwab, R., Carrington, J.C., and Weigel, D. (2003). Control of leaf morphogenesis by microRNAs. Nature **425:** 257-263.

Pillitteri, L.J., Sloan, D.B., Bogenschutz, N.L., and Torii, K.U. (2007). Termination of asymmetric cell division and differentiation of stomata. Nature **445:** 501-505.

Rai, M.I., Wang, X., Thibault, D.M., Kim, H.J., Bombyk, M.M., Binder, B.M., Shakeel, S.N., and Schaller, G.E. (2015). The ARGOS gene family functions in a negative feedback loop to desensitize plants to ethylene. BMC Plant Biol **15:** 157.

Reinhardt, D., Pesce, E.R., Stieger, P., Mandel, T., Baltensperger, K., Bennett, M., Traas, J., Friml, J., and Kuhlemeier, C. (2003). Regulation of phyllotaxis by polar auxin transport. Nature **426:** 255-260.

Rodriguez, R.E., Debernardi, J.M., and Palatnik, J.F. (2014). Morphogenesis of simple leaves: regulation of leaf size and shape. Wiley Interdiscip. Rev.: Dev. Biol. **3:** 41-57.

Rodriguez, R.E., Mecchia, M.A., Debernardi, J.M., Schommer, C., Weigel, D., and Palatnik, J.F. (2010). Control of cell proliferation in Arabidopsis thaliana by microRNA miR396. Development **137:** 103-112.

Sack, L., Scoffoni, C., McKown, A.D., Frole, K., Rawls, M., Havran, J.C., Tran, H., and Tran, T. (2012). Developmentally based scaling of leaf venation architecture explains global ecological patterns. Nat. Commun. **3:** 837.

Sakamoto, T., Kamiya, N., Ueguchi-Tanaka, M., Iwahori, S., and Matsuoka, M. (2001). KNOX homeodomain protein directly suppresses the expression of a gibberellin biosynthetic gene in the tobacco shoot apical meristem. Genes Dev **15:** 581-590.

Scarpella, E., Francis, P., and Berleth, T. (2004). Stage-specific markers define early steps of procambium development in Arabidopsis leaves and correlate termination of vein formation with mesophyll differentiation. Development **131:** 3445-3455.

Scarpella, E., Marcos, D., Friml, J., and Berleth, T. (2006). Control of leaf vascular patterning by polar auxin transport. Genes Dev. **20:** 1015-1027.

Schmid, M., Uhlenhaut, N.H., Godard, F., Demar, M., Bressan, R., Weigel, D., and Lohmann, J.U. (2003). Dissection of floral induction pathways using global expression analysis. Development **130:** 6001-6012.

Schommer, C., Palatnik, J.F., Aggarwal, P., Chételat, A., Cubas, P., Farmer, E.E., Nath, U., and Weigel, D. (2008). Control of jasmonate biosynthesis and senescence by miR319 targets. PLoS Biol. **6:** e230.

Schoof, H., Lenhard, M., Haecker, A., Mayer, K.F., Jurgens, G., and Laux, T. (2000). The stem cell population of Arabidopsis shoot meristems in maintained by a regulatory loop between the CLAVATA and WUSCHEL genes. Cell **100:** 635-644.

Spartz, A.K., Ren, H., Park, M.Y., Grandt, K.N., Lee, S.H., Murphy, A.S., Sussman, M.R., Overvoorde, P.J., and Gray, W.M. (2014). SAUR inhibition of PP2C-D phosphatases activates plasma membrane H+-ATPases to promote cell expansion in Arabidopsis. The Plant Cell Online **26:** 2129-2142.

Spartz, A.K., Lee, S.H., Wenger, J.P., Gonzalez, N., Itoh, H., Inzé, D., Peer, W.A., Murphy, A.S., Overvoorde, P.J., and Gray, W.M. (2012). The *SAUR19* subfamily of *SMALL AUXIN UP RNA* genes promote cell expansion. Plant J. **70:** 978-990.

Stamm, P., and Kumar, P.P. (2013). Auxin and gibberellin responsive Arabidopsis SMALL AUXIN UP RNA36 regulates hypocotyl elongation in the light. Plant Cell Rep. **32:** 759-769.

Stransfeld, L., Eriksson, S., Adamski, N.M., Breuninger, H., and Lenhard, M. (2010). *KLUH/CYP78A5* promotes organ growth without affecting the size of the early primordium. Plant Signal. Behav. **5:** 982-984.

Sugimoto-Shirasu, K., and Roberts, K. (2003). "Big it up": endoreduplication and cell-size control in plants. Curr. Opin. Plant Biol. **6:** 544-553.

Sullivan, M., and Morgan, D.O. (2007). Finishing mitosis, one step at a time. Nat. Rev. Mol. Cell Biol. **8:** 894-903.

Takada, S., Hibara, K.-i., Ishida, T., and Tasaka, M. (2001). The CUP-SHAPED COTYLEDON1 gene of Arabidopsis regulates shoot apical meristem formation. Development **128:** 1127-1135.

Takahashi, K., Hayashi, K., and Kinoshita, T. (2012). Auxin activates the plasma membrane H+-ATPase by phosphorylation during hypocotyl elongation in Arabidopsis. Plant Physiol **159:** 632-641.

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. Nature **408:** 796-815.

Torii, K.U. (2015). Stomatal differentiation: the beginning and the end. Curr. Opin. Plant Biol. **28:** 16-22.

Tsukaya, H. (2013). Does ploidy level directly control cell size? Counterevidence from Arabidopsis genetics.

Veit, B. (2009). Hormone mediated regulation of the shoot apical meristem. Plant Mol Biol **69:** 397-408.

Vroemen, C.W., Mordhorst, A.P., Albrecht, C., Kwaaitaal, M.A., and de Vries, S.C. (2003). The CUP-SHAPED COTYLEDON3 gene is required for boundary and shoot meristem formation in Arabidopsis. Plant Cell **15:** 1563-1577.

Wang, J.-W., Czech, B., and Weigel, D. (2009). miR156-regulated SPL transcription factors define an endogenous flowering pathway in Arabidopsis thaliana. Cell **138:** 738-749.

Wang, J.-W., Schwab, R., Czech, B., Mica, E., and Weigel, D. (2008). Dual effects of miR156-targeted SPL genes and CYP78A5/KLUH on plastochron length and organ size in Arabidopsis thaliana. The Plant Cell **20:** 1231-1243.

Weinl, C., Marquardt, S., Kuijt, S.J., Nowack, M.K., Jakoby, M.J., Hülskamp, M., and Schnittger, A. (2005). Novel functions of plant cyclin-dependent kinase inhibitors, ICK1/KRP1, can act non-cell-autonomously and inhibit entry into mitosis. The Plant Cell **17:** 1704-1722.

Werner, T., Motyka, V., Laucou, V., Smets, R., Van Onckelen, H., and Schmulling, T. (2003). Cytokinin-deficient transgenic Arabidopsis plants show multiple developmental alterations indicating opposite functions of cytokinins in the regulation of shoot and root meristem activity. Plant Cell **15:** 2532-2550.

White, D.W.R. (2006). *PEAPOD* regulates lamina size and curvature in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **103:** 13238-13243.

Yadav, R.K., and Reddy, G.V. (2011). WUSCHEL-mediated cellular feedback network imparts robustness to stem cell homeostasis. Plant Signal Behav **6:** 544-546.

Zhao, Y. (2010). Auxin biosynthesis and its role in plant development. Annu. Rev. Plant Biol. **61:** 49.

Zotz, G., Wilhelm, K., and Becker, A. (2011). Heteroblasty—a review. The Botanical Review **77:** 109-151.

Chapter 2

Leaf growth promoting genes

Youn-Jeong Nam[1,2], Dirk Inzé[1,2*] & Nathalie Gonzalez[1,2*]

[1]Department of Plant Systems Biology, VIB B-9052 Gent, Belgium, [2]Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Gent, Belgium
* These authors contributed equally to this work.

Contributions: Y.J.N. was the main author of this chapter. D.I. and N.G. contributed to the writing of this chapter.

Chapter 2. Leaf growth promoting genes

2.1 Leaf growth promoting genes

As described in Chapter 1, many genes regulating the different phases of leaf development/growth have been described as leaf growth promoting genes, since increased leaf size is observed when their expression is altered (Gonzalez et al., 2009; Rojas et al., 2010). In Table 2.1, a summary of the different leaf growth promoting genes described in Arabidopsis is shown. These genes can be classified in function of their biological functions. For example, numerous genes involved in the regulation of transcription including transcription factors and transcriptional co-activators or co-inhibitors have been identified. Some genes are involved in post transcriptional regulation via miRNAs. Other leaf growth promoting genes were shown to have a role in protein synthesis and modification. Several growth promoting genes encode proteins involved in the regulation of hormone biosynthesis or signalling including GA, CK, brassinosteroid and auxin signalling. In conclusion, genes encoding proteins or miRNAs involved in various biological processes can stimulate plant growth.

Most research on leaf growth regulation has been performed using Arabidopsis as a model system. In some cases, the growth promoting genes identified in Arabidopsis also show similar positive effect on growth in crops including monocots. For instance, in both Arabidopsis and tomato, high expression level of *HERCULES1* (*HRC1*), an AT-hook family transcription factor, leads to increased organ size (Century et al., 2008). In rice, the level of EXP4, one of the EXPANSIN proteins, is positively correlated with seedling growth rate (Choi et al., 2003). Overexpression of *ARGOS* promotes plant and organ growth in Arabidopsis (Hu et al., 2003; Hu et al., 2006) but also maize (Guo et al., 2014). To extend our knowledge on the regulation of leaf growth and the transferability of phenotypes, evaluating the effects of various growth promoting genes in other genotypes such as another accession or species is of great interest to researchers and potentially to breeders.

Table 2.1 Genes involved in leaf development and their phenotype when the expression is altered (OE; over-exprssion, LOF; loss-of-function)*;gene that identified in rice

Genes	Alteration	Phenotype	References
Transcriptional regulation			
SWP	OE	Leaves with increased number of small cells	Autran et al., 2002
quadruple-DELLA	LOF	Bigger leaves in early developmental stage	Achard et al., 2009
MED25	LOF	Large organs, with larger and slightly increased numbers of cells	Xu and Li, 2011
ANT	OE	Enlarged embryonic and all shoot organs	Mizukami and Fischer, 2000
ARF2	LOF	Large, dark green rosette leaves, large organ size	Okushima et al., 2005
ATAF2	OE	Increased biomass	Delessert et al., 2005
ATHB16	LOF	Increased size of the leaves	Wang et al., 2003
GIF1/AN3	OE	Increased leaf size	Horiguchi et al., 2005
GRF1, 2, 3, 5	OE	Leaves and cotyledons growth promotion	Kim et al., 2003; Horiguchi et al., 2005
HRC1	OE	Increases plant organ size and yield	Century et al., 2008
NAC1	OE	Plants with more abundant roots, larger leaves and thicker stems	Xie et al., 2000
OBP2	LOF	Increased leaf size	Skirycz et al., 2006
PPD	LOF	larger leaf with dome-shape	White, 2006
RON2	LOF	Wide and serrated vegetative leaf lamina	Cnops et al., 2004
ZHD5	OE	Larger organs containing larger cells	Hong et al., 2011
microRNAs			
miR396	LOF	Larger leaves with more cells	Rodriguez et al., 2010
miR319/JAW	OE	Larger leaves with morphology changes	Palatnik et al., 2003; Schommer et al., 2008
miR156	OE	Increase in total leaf number on main and side shoots	Scchwab et al., 2005
Protein synthesis and modification			
DA1 and *DAR1*	LOF	Larger seeds, embryos, cotyledon,leaves and flowers, thicker stems	Li et al., 2008
DA2	LOF	Larger seeds and organsize	Xia et al., 2013
RPT2a	LOF	Enlarged leaves caused by increased cell size	Sonoda et al., 2009
EBP1	OE	Larger leaf size with more cells and larger cells	Horváth et al., 2006
Cell cycle machinery			
CYCD2	OE	Increased rate of leaf initiation and accelerated development	Cockcroft et al., 2000
CYCD3	OE	Leaves with more but smaller cells	Dewitte et al., 2003
ABAP1	LOF	Larger leaves with more cells	Masuda et al, 2008
CDC27a	OE	Increased growth rate and organ size	Rojas et al., 2009
APC10	OE	Enhanced leaf size	Eloy et al., 2011
SAMBA	LOF	Larger seeds, leaves, and roots	Eloy et al., 2012
CCS52A	milder OE	Increased organ sizes	Baloban et al., 2013

Continued

Genes	Alteration	Phenotype	References
Hormone regulation			
Brassinosteroid			
ARL (ARGOS-LIKE)	OE	Increased leaf size	Hu et al., 2006
BEN1	LOF	Increased leaf size	Yuan et al., 2007
BRI1	OE	Increased leaf size	Wang et al., 2001
DWF4	OE	Increased leaf size	Choe et al., 2001
EXORDIUM	OE	Increased leaf size	Coll-Garcia et al., 2004
TTL	LOF	Increased leaf size	Nam and Li, 2004
DAS5	OE	Increased plant fresh weight	Chory et al., 2004
Auxin			
SAUR19	stabilized OE with fusion proteins	Increased hypocotyl and leaf size	Spartz et al., 2012
SAUR36	LOF	Increased leaf size with larger cells	Hou et al., 2013
AVP1	OE	Increment in the number and size of rosette leaves and in root size	Li et al., 2005
ARGOS	OE	Larger organs	Hu et al., 2003
Gibberellin			
GA20-oxidase	OE	Larger leaves, longer petiole, early flowering	Coles et al., 1999; Huang et al., 1998
Cytokinin			
HOG1	LOF	Increments in leaf size and seed yield	Godge et al., 2008
IPT	OE	Increased leaf biomass	Rupp et al., 1999
Cell wall extension			
EXP3	OE	Increased leaf size	Kwon et al., 2008
EXP10	OE	Increased leaf size	Cho et al., 2000
Photosynthsis			
*Phosphoenolpyruvate carboxylase /Pyruvate orthophosphate dikinase**	OE	Enhanced stomatal conductance, Increased photosynthetic capacity and tiller number	Matsuoka et al., 2001, Ku et al., 2007
Cytochrome c6	OE	Higher content of photosynthetic metabolites Increased leaf and root growth	Chida et al., 2007
Rubisco activase	Gene shuffling	More siliques and enhanced vegetative growth	Kurek et al., 2007
Glycolate dehydrogenase/Glyoxylate carboligase/Tartronic semialdehyde reductase	OE	Improvement of carboxilation/oxygenation ratio, greater biomass and an increase in photosynthesis	Kebeish et al., 2007
Other signaling pathways			
CLE26	OE	Larger leaves	Strabala et al., 2006
KLU	OE	Larger leaves with more cells	Anastasiou et al., 2007
SRF4	OE	Larger leaves	Eyüboglu et al., 2007
Other			
KAT2	LOF	Larger rosette area and higher dryweight	Footitt et al., 2007
GRA	unknown	Larger leaf with more cells	Horiguchi et al., 2009

Among the leaf growth promoting genes described in table 2.1, three genes *GA20ox1*, *PPD*, and *DA1* encoding gibberellin biosynthesis enzyme, transcription regulator, and ubiquitin receptor, respectively, were chosen for further analysis in this thesis. These three leaf growth promoting genes were selected since they have different biological functions and they are involved in the regulation of different processes during leaf development. Furthermore, in the System Biology of Yield group, where I did my PHD, these three genes have been investigated by other lab members, therefore, it was the best choice to select these genes to work with. The details about these three genes will be discussed below.

2.2. Gibberellin and its rate limiting biosynthesis enzyme, GA20ox1

2.2.1 Importance of plant hormones for plant growth

Plant hormones are chemicals produced in one part of the plant and having their physiological effect on a target tissue. Leaves, stem, root, flowers, seeds, and fruits all produce hormones. Most plant hormones are functional at extremely low concentrations. When hormones reach the target tissue they can have a direct effect causing a rapid metabolic response or they involve the intervention of a secondary messenger within the target cells that can affect transcription of target genes.

Auxins, cytokinins, gibberellins, abscisic acid, and ethylene are the best known plant hormones and have roles in many aspects of plant development (Davies, 2012). Auxins were the first class of plant hormones identified that positively influence bud formation and root initiation (Zhao, 2010). Cytokinins delay senescence and activate dormant buds (Ha et al., 2012). Abscisic acid (ABA) has roles in maintenance of dormancy in seeds and buds and stimulation of stomatal closure (Finkelstein, 2013). Ethylene is the only plant hormone that is a gas and it is involved in the regulation of leaf abscission and causes fruit ripening (Chang and Bleecker, 2004). Jasmonates (JA) are involved in response to environmental stresses and control of seed germination (Carvalhais et al., 2013). Salicylic acid (SA) activates genes involved in plant defence mechanisms (Rivas-San Vicente and Plasencia, 2011). Brassinosteroids are plant steroids and they regulate vascular differentiation and flowering (Fariduddin et al., 2014).

Some of these hormones are involved in regulating plant growth. For instance, auxins regulate cell division and expansion (Zhao, 2010). Cytokinins have a positive effect on cell division (Ha et al., 2012). Cytokinins work synergistically with auxin in the control of tissue and organ differentiation. Brassinosteroids promote both cell elongation and division (Fariduddin et al., 2014).

Gibberellins (GAs) also play crucial roles in plant development including germination, stem elongation, leaf, flower and fruit development and affect both cell division and expansion (Olszewski et al., 2002; Hedden, 2003; Ueguchi-Tanaka et al., 2007; Schwechheimer and Willige,

2009). GAs are also important signalling molecules for regulating growth in response to environmental changes (Yamaguchi, 2008; Harberd et al., 2009; Claeys et al., 2014).

Because of the importance of plant hormones, hormone signalling has become a good target for genetic modifications to understand the underlying mechanisms (Hedden and Phillips, 2000; Santner and Estelle, 2009). For example, in 1960's, the dramatic increases in grain yield of wheat and rice during the "Green Revolution" were enabled by introducing a semi-dwarf growth habit caused by mutation in genes, *Reduced height*; *Rht* in wheat and *semidwarf*; *sd1* in rice, involved in GA signal transduction and GA biosynthesis, respectively (Peng et al., 1999; Monna et al., 2002; Sasaki et al., 2002; Spielmeyer et al., 2002).

2.2.2 GA metabolism

GAs are diterpene hormones and bioactive GA forms are synthesized through a complex pathway (Figure 2.1). Most of the genes encoding enzymes involved in GA biosynthesis and catabolism have been identified via biochemical, genetic, and genomic approaches (Yamaguchi, 2008). Geranylgeranyldiphostphate (GGDP), a common C_{20} precursor for diterpenoids, is at the origin of GAs. Three different classes of enzymes are involved in converting GGDP into the first GA intermediate: TERPENE SYNTHASES (TPSs), CYTOCHROME 450 MONOOXYGENASES (P450s), and 2-OXOGLUTARATE-DEPENDENT DIOXYGENASES (2ODDs). Two TPS, ENT-COPALYL DIPHOSPHATE SYNTHASES (CPS) and ENT-KAURENE SYNTHASE (KS) are located in the plastids (Sun and Kamiya, 1994; Aach et al., 1997; Song et al., 1997; Helliwell et al., 2001) and convert GGDP into CDP and then into ent-Kaurene. Two P450s, ENT-KAURENE OXIDASE (KO) and ENT-KAURENOIC ACID OXIDASE (KAO), convert ent-Kaurene into ent-kaurenoic acid and then into GA_{12}. These two P450s are located in the outer membrane of the plastids and the endoplasmic reticulum, respectively (Helliwell et al., 2001). GA_{12} is converted into the bioactive forms of GAs, GA_1 and GA_4, through several steps of oxidation on C-20 and C-3 by GA20OXIDASE (GA20ox) and GA3OXIDASE (GA3ox), respectively. Both enzymes are soluble and located in the cytosol (Yamaguchi, 2008). There are two different pathways to synthesize the bioactive GA forms from GA_{12} which are the non-13-hydroxyl pathway and the 13-hydroxyl pathway in which the first intermediate is GA_{53} formed by GA13OXIDASE (GA13ox) using GA_{12} as substrate. The two bioactive forms, GA_1 and GA_3, originate from the 13-hydroxyl pathway and the two other bioactive forms, GA_4 and GA_7 derive from the non-13-hydroxyl pathway (Sun, 2011).

GA20ox*,* a rate limiting enzyme in the GA biosynthesis pathway, catalyses consecutive steps of oxidation in the late part of the formation of active GAs (Huang et al., 1998; Coles et al., 1999). GA20ox uses various intermediates as substrates, including GA_{12}, GA_{53}, GA_{15}, GA_{44}, GA_{24}, and GA_{19}, through successive oxidative reactions to form as a final step GA_9 and/or GA_{20}, two substrates of GA3ox, that are then converted into bioactive GAs (Hedden and Thomas, 2012).

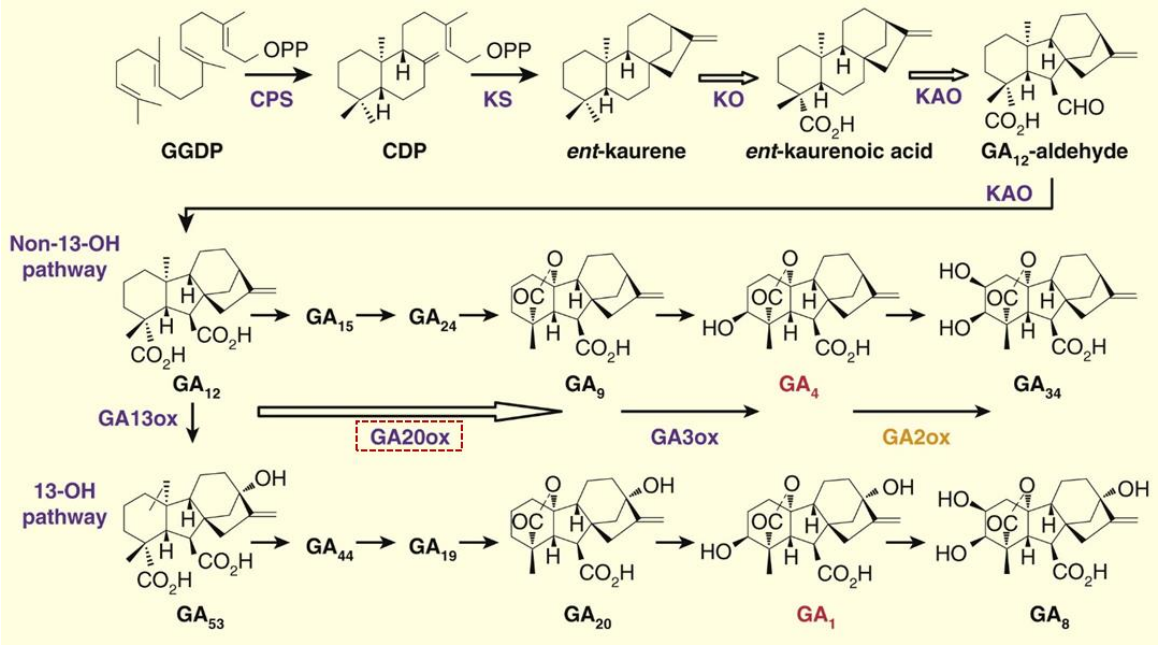


Figure 2.1 GA biosynthesis and deactivation pathway in plants. The simplified GA biosynthesis and deactivation pathway shows the different GAs (intermediates, bioactive and inactive forms) and enzymes involved. Active GAs are labelled in red, GA biosynthesis enzymes in purple, and the deactivation enzyme in orange. The red dashed box highlights GA20ox, a rate limiting enzyme in the GA biosynthesis pathway. The black arrows indicate a single-step reaction. The unfilled arrows indicate a multiple-step reaction (from Sun, 2011).

Deactivation of hormones is an important step to fine-tune their concentrations. GA2ox, a member of the 2ODDs, is a major enzyme converting bioactive GA forms (and/or their precursors) into bio-inactive forms (Figure 2.1). Methylation of GAs, triggered by GIBBERELLIN METHYLTRANSFERASE (GAMT) in Arabidopsis also leads to deactivation of GAs (Varbanova et al., 2007).

As illustrated above, the regulation of GA levels involves multiple enzymes and is also quite complex. For example, 2ODDs enzymes acting in the late stages of the GA biosynthesis and deactivation are important for hormone homeostasis via feedback mechanisms in which the outputs of a process are routed back as inputs to regulate this same process (Hedden and Phillips, 2000; Yamaguchi, 2008). In addition, treatment with GA biosynthesis inhibitor results in up-regulation of the GA biosynthesis genes (*GA20ox* and *GA3ox*) and down-regulation of GA deactivation gene (*GA2ox*) allowing to maintain homeostasis.

2.2.3 The GA-GID1-DELLA signalling pathway

The molecular mechanisms downstream of GA have been well characterized via genetic, biochemical, and structural studies (Figure 2.2). The first identified GA receptor was GA INSENSITIVE DWARF 1 (GID1) from rice. GID1 is localized both in the cytoplasm and in the nucleus (Ueguchi-Tanaka et al., 2005; Willige et al., 2007). Arabidopsis genome contains three *GID1*

(*AtGID1*) orthologs: *GID1A*, *GID1B*, and *GID1C*. AtGIDs are soluble GA receptors having high affinity only with bioactive GAs but not with other GA derivatives (Nakajima et al., 2006; Ueguchi-Tanaka et al., 2007). Upon GA binding, conformational changes occur in GID1 allowing interaction with the GRAS (GIBBERELLIC-ACID INSENSITIVE (GAI), REPRESSOR of GAI (RGA) and SCARECROW (SCR)) domain of the DELLA protein. DELLA proteins are members of the plant-specific GRAS family which are nuclear transcription regulators that suppress the GA response and therefore inhibit plant growth (Peng et al., 1997; Silverstone et al., 1998).The GA-GID1-DELLA complex is recognized by a specific ubiquitin E3 ligase complex, the $SCF^{SLY1/GID2}$. $SCF^{SLY1/GID2}$ promotes a rapid degradation of DELLAs through the ubiquitin-proteasome machinery thereby repressing their growth-restraining effects (McGinnis et al., 2003; Sasaki et al., 2003; Dill et al., 2004; Fu et al., 2004). In plants, DELLAs are highly conserved in different species, such as Arabidopsis, rice, barley, maize, and wheat (Peng et al., 1997; Peng et al., 1999; Ikeda et al., 2001; Chandler et al., 2002). Five DELLA proteins are present in Arabidopsis: REPRESSOR OF GA1-3(RGA); RGA-LIKE1 (RGL1); RGL2 and RGL3; GA-INSENSITIVE (GAI). The different DELLA proteins not only have redundant functions but also play distinct roles in the GA response (Dill et al., 2001; King et al., 2001; Lee et al., 2002; Tyler et al., 2004) implying a dynamic regulation of GA-mediated growth response. Besides GA-dependent proteolysis, DELLA activity can also be controlled by other mechanisms such as glycosylation and phosphorylation. SPINDLY (SPY), an *O*-LINKED GlcNAc TRANSFERASE (OGT), is known to activate DELLAs, probably by glycosylation although it is not clear whether DELLAs are direct targets of SPY (Silverstone et al., 2007). In rice, EARLIER FLOWERING1 (EL1), a casein kinase, has been suggested to increase DELLA stability via phosphorylation (Dai and Xue, 2010).

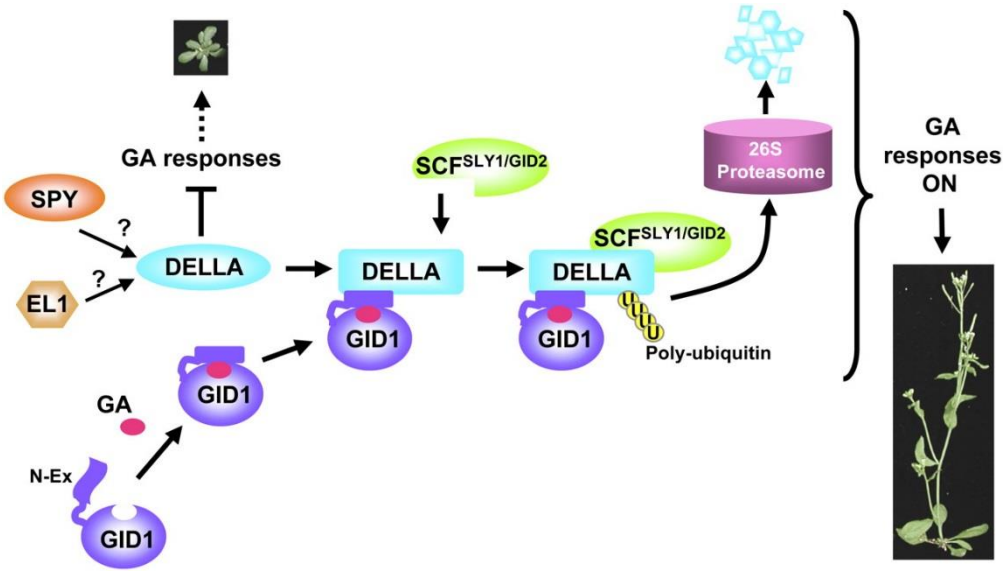


Figure 2.2 Model of GA signalling in plants. To degrade DELLA, a repressor of GA response, binding of bioactive GA to GID1 induces a conformational change in the N-terminal extension of GID1 allowing DELLA binding. This assembly promotes a conformational transition in the GRAS domain of the DELLA protein for $SCF^{SLY1/GID2}$ recognition. DELLA proteins will then be poly-ubiquitinated and degraded via the ubiquitin-proteasome pathway. SPY may activate DELLA by GlcNAc-modification, whereas EL1 (a casein kinase in rice) may phosphorylate and activate DELLA (from Sun, 2011).

2.2.4 Mechanism of GA-GID1-DELLA regulation of plant growth

The GA-GID1-DELLA network is an important regulatory component that controls plant growth and development through integration of internal signals from other hormone pathways (auxin, abscisic acid, cytokinin, ethylene and brassinosteroid) and external factors such as light condition (Figure 2.3). Since DELLAs do not have DNA-binding domain and the promoters of DELLA targets do not contain any conserved responsive *cis*-elements, it has been suggested that DELLAs regulate target genes expression by interacting with different transcription factors (Zentella et al., 2007). For example, DELLA interacts with PHYTOCHROME INTERACTING FACTORs (PIFs), members of subfamily 15 of bHLH transcription factors in Arabidopsis. The DELLA-PIF interaction inhibits PIF-mediated hypocotyl elongation by repressing the transcription of PIF's target genes (de Lucas et al., 2008; Feng et al., 2008).

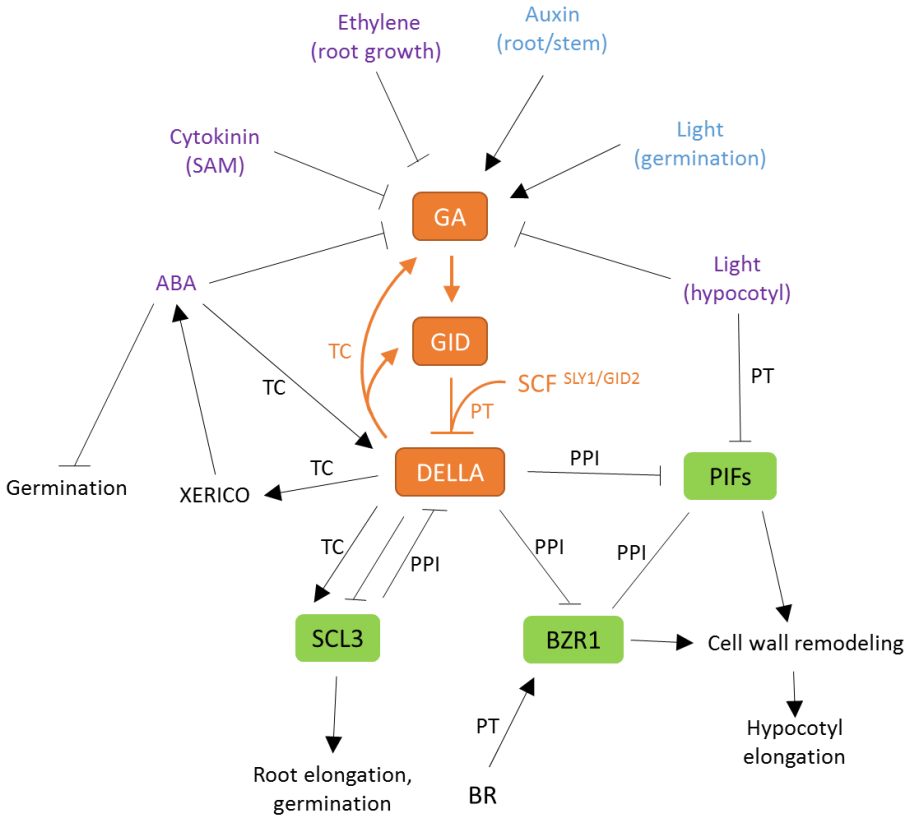


Figure 2.3 The GA–GID1–DELLA signalling module and other internal and external factors. The GA–GID1–DELLA regulatory module is highlighted in orange. Signals that promote bioactive GA accumulation are labelled in blue, whereas signals that reduce GA levels are highlighted in purple. DELLA interacts directly with multiple regulatory proteins (PIFs, SCL3 and BZR1; highlighted in green) to mediate crosstalk between GA and other pathways (light and JA signalling, and root and fruit patterning pathways). Activation or inhibition of factors involved in this network could be via different modes of action: TC, transcriptional regulation; PT, post-transcriptional regulation; PPT, protein-protein interaction. SAM, shoot apical meristem; ABA, abscisic acid; BR, Brassinosteroid (modified from Sun, 2011).

When GA is absent, DELLA proteins activate or inhibit the expression of several target genes (Claeys et al., 2014). DELLAs induce the expression of downstream negative components of GA response such as putative transcription factors/regulators or RING type ubiquitin E2/E3 ligases. For example, *XERICO* encoding an E3 ligase, induces the accumulation of abscisic acid (ABA) that antagonizes GA response (Ko et al., 2006; Zentella et al., 2007). Several DELLA-induced target genes encode GA biosynthesis enzymes or GA receptors, showing that DELLAs are also involved in maintaining GA homeostasis via feedback regulation (Zentella et al., 2007). *SCARECROW-LIKE3* (*SCL3*), a DELLA-induced target gene, plays a role as positive regulator of GA signalling and is an inhibitor of DELLA proteins (Heo et al., 2011; Zhang et al., 2011). Although SCL3 is also a member of the GRAS family, it does not contain GA-responsive DELLA domain. Instead, SCL3 directly interacts with DELLA to regulate target gene expression (Zhang et al., 2011). Interestingly, SCL3-DELLA interaction functions not only downstream GA signalling but also maintains GA homeostasis via transcriptional regulation of GA biosynthetic genes (Zhang et al., 2011).

2.2.5 Overexpression of *GA20ox1* leads to increased leaf size

Overexpression of *GA20ox1* in Arabidopsis Col-0 background has been shown to enhance growth since the transgenic plants are taller, with larger leaves, and flower earlier (Figure 2.4) (Huang et al., 1998; Coles et al., 1999; Gonzalez et al., 2010). The increased leaf size in these plants overexpressing *GA20ox1* results from an increase in cell number and cell size (Gonzalez et al., 2010). Also in other plants such as potato (Carrera et al., 2000), poplar (Eriksson et al., 2000; Israelsson et al., 2005), rice (Oikawa et al., 2004), tobacco (Vidal et al., 2001; Biemelt et al., 2004) and maize (Nelissen et al., 2012) growth is enhanced upon GA20ox overexpression. Enhanced plant growth in these overexpressing plants results from an increase in GA levels (Huang et al., 1998; Coles et al., 1999; Gonzalez et al., 2010; Nelissen et al., 2012). Arabidopsis genome contains five *GA20ox* paralogs (Hedden and Phillips, 2000). Mutant analysis of *GA20ox* genes suggested a complex regulatory relationship between these five paralogs with *GA20ox1*, *-2*, and *-3* being the dominant paralogs and *GA20ox4* and *-5* having very minor roles (Plackett et al., 2012). The triple mutant of *GA20ox1*, *-2*, and *-3* genes presents a severe dwarf phenotype and is almost infertile. In addition, the functions of *GA20ox1*, *-2*, and *-3* are almost completely redundant (Plackett et al., 2012). In Arabidopsis, however, plants overexpressing *GA3ox,* also encoding GA biosynthesis enzymes, do not show enhanced growth phenotype in comparison to *GA20ox1*^{OE} transgenics (Phillips, 2010).

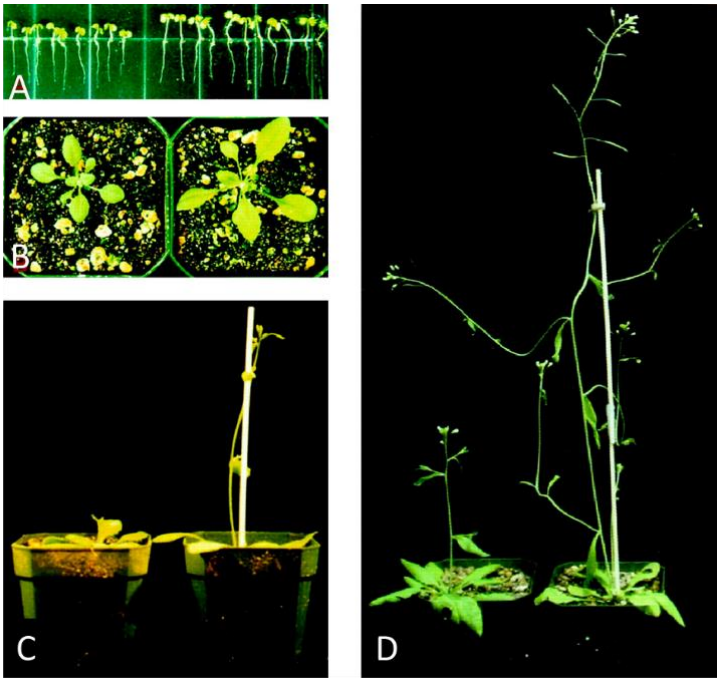


Figure 2.4 Phenotype of *GA20ox1*^{OE} in Col-0 background. The pictures compare the growth of transgenic (right) and wild type (left) plants at 1 week (A), 3 weeks (B), 3.5 weeks (C), and 4 weeks (D) (from Huang et al., 1998).

2.3 PEAPOD, transcription factors that negatively regulate meristemoid division

The timing of meristemoid division can be modified by modulating the expression of the *PEAPOD* (*PPD*) genes encoding transcriptional regulators. Arabidopsis genome encodes two PPD proteins, PPD1 and PPD2, which share 84 % identity. PPD proteins belong to the plant-specific TIFY transcription factor family (Figure 2.5) (Bai et al., 2011; Cuellar Perez et al., 2014). Arabidopsis contains 18 TIFY proteins which are grouped in two classes according to the presence or absence of a C2C2-GATA domain, class l and class ll, respectively (Cuellar Perez et al., 2014). PPD proteins belong to class ll that contains also 12 JASMONATE ZIM DOMAIN (JAZ) proteins and TIFY8. TIFY proteins contain a highly conserved TIFY domain which is formed of about 28 amino acids and a core motif TIF[F/Y]XG. This domain (also called ZIM (ZINC-FINGER PROTEIN EXPRESSED IN INFLORESCENCE MERISTEM) domain) mediates interactions between TIFY proteins (homo- and heteromeric dimerization) and other specific transcription factors (Chini et al., 2009). JAZ proteins interact with NOVEL INTERACTOR JAZ (NINJA), an adaptor protein, through their ZIM domain to recruit the Groucho/Tup1-type corepressor TOPLESS (TPL) and both NINJA and TPL act as transcriptional repressors of the jasmonate signalling (Pauwels et al., 2010). Most of the class ll TIFY proteins (except TIFY 8) have a C-terminal Jas domain known to mediate either binding to a bHLH MYC factor or the F-box protein COI1 depending on the absence or the presence of JA (Pauwels et al., 2010). The PPD proteins have a divergent C-terminal Jas domain and an additional N-terminal PPD-domain compared to other JAZ proteins (Figure 2.5) (Bai et al., 2011).

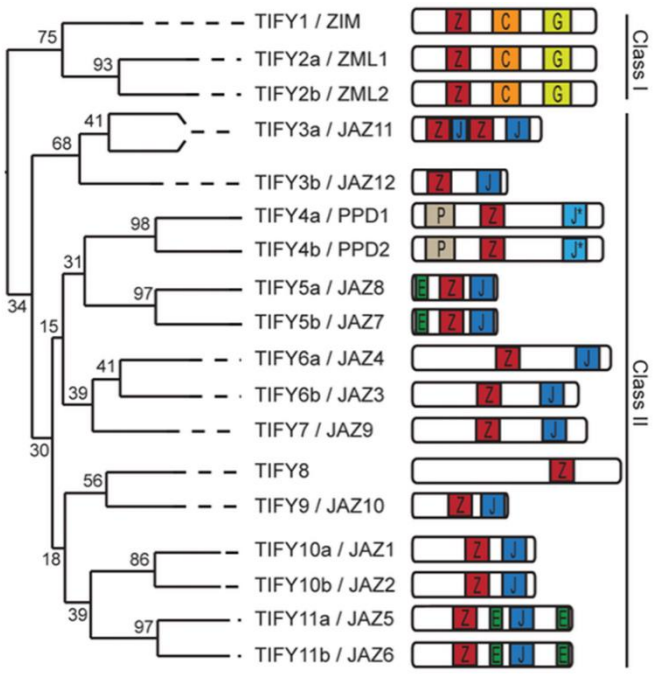


Figure 2.5 The TIFY proteins family in Arabidopsis. Phylogenetic tree of the Arabidopsis TIFY family members based on the ZIM domain (Z) protein sequence. The TIFY family can be divided in two classes according to the presence or absence of a C2C2-GATA domain. C: CONSTANS, CO-like, and TOC1 (CCT) domain; G: C2C2-GATA Zn-finger; P: PEAPOD domain; J: Jas domain; J* Jas-like domain; E: EAR domain. (from Cuellar Perez et al., 2014).

In Arabidopsis, deletion or down-regulation of the expression of the *PPD* genes causes the formation of larger leaves with a dome-shape phenotype as a result of prolonged meristemoid division (Figure 2.6) (White, 2006; Gonzalez et al., 2015). This dome-shape is caused by an increased area with almost no changes in the perimeter of the leaf blade. Not only leaf area is affected by the deletion of PPD, but also siliques are shorter, flattened, and wider. In addition, in the cotyledons a more extensive vascular network is observed (White, 2006). In contrast, overexpression of *PPD* leads to the formation of smaller and flat leaves (White, 2006).

A recent study discovered by tandem affinity purification (TAP) that PPD2 interacts, through its PPD domain, with KINASE INDUCIBLE DOMAIN INTERACTING (KIX) 8 and 9 proteins which might function as adaptor proteins for the corepressor TPL (Gonzalez et al., 2015). KIX proteins are well documented in non-plant species and function as transcriptional coactivators (Radhakrishnan et al., 1997; Yang et al., 2006). In Arabidopsis, 11 KIX proteins exist and a KIX domain specific to these proteins is located in their N-terminal region (Thakur et al., 2013). Interestingly, the *kix8 kix9* double mutant phenocopies the leaf phenotype of the *amiPPD* mutant in Col-0 background (Gonzalez et al., 2015). The double mutant *kix8 kix9* does not only lead to an increased leaf area with a dome-shape but also increases cell number especially the number of small cells surrounding the stomata. However, *kix8* and *kix9* single mutants do not show a clear phenotype suggesting a redundancy of the function of KIX8 and KIX9.

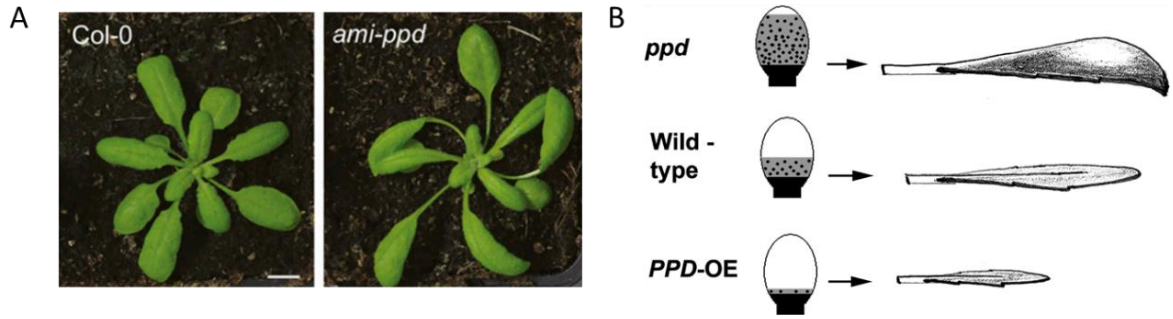


Figure 2.6 Effect of alteration of *PPD* expression in leaf. (A) Phenotype of *amiPPD* in Col-0 background. Wild type (left) and *amiPPD* (right) plants grown for 25 DAS in soil (from Gonzalez et al., 2015). Bar = 1 cm (B) Schematic illustration showing the effect of *PPD* expression levels on meristemoid division, leaf size, and curvature. The black area represents the cell proliferation zone. The meristemoid division zone is shown in gray color; regions without cell division are white; black dots represent the meristemoid cells (from White, 2006).

Down-regulation of *PPD2* leads to the differential expression of a specific set of genes (Gonzalez et al., 2015). Two main categories are represented in the DE genes that are mainly up regulated: stomatal lineage related genes and proliferation-specific genes. Tandem chromatin affinity purification followed by sequencing identified DNA sequences bound by PPD2. The target genes identified are mainly part of the functional categories “regulation of transcription” and “hormone metabolism”. Among the genes bound by PPD2, thirteen were found to overlap with the set of 36 differentially expressed genes.

Two cell cycle-related genes *CYCD3;2* and *CYCD3;3* were found to be targets of PPD2 (Gonzalez et al., 2015). PPD2 binds to the promoter region of these two genes and down-regulation of *PPD2* causes their up-regulation. Three D3-type cyclins are encoded in the Arabidopsis genome and their combined mutation induces early cell cycle arrest resulting in decreased cell number in the leaf (Dewitte et al., 2007). Interestingly, this triple mutant shows reduced meristemoid division rate and lower frequency of formation of satellite meristemoids in the leaf (Elsner et al., 2012). Moreover*, CYCD3;2* is found as one of direct targets of SPEECHLESS (SPCH), a positive regulator of meristemoid division, and induction of SPCH leads to up-regulation of *CYCD3;2* (Lau et al., 2014). It was proposed that PPD2 is important for the regulation of meristemoid activity via the transcriptional repression of *CYCD3* genes.

2.4 DA1 and its regulation of organ growth

2.4.1 Ubiquitination mediated control of plant development

In plants, regulation of protein degradation by the ubiquitin/26S proteasome contributes substantially to development by modulating a variety of processes such as embryogenesis, hormone signalling and senescence (Brooks, 1973). Ubiquitination corresponding to the transfer of an ubiquitin protein to a substrate protein occurs through the action of three enzymes:

ubiquitin activating enzyme (E1), ubiquitin conjugating enzyme (E2), and ubiquitin protein ligase (E3). In most cases, ubiquitination of a protein corresponds to a signal for degradation through the proteasome, changes its cellular location, controls its specific activity, or induces or inhibits interaction with other proteins. In Arabidopsis, about 1400 genes encode components of the ubiquitin/26S proteasome pathway (Smalle and Vierstra, 2004). Among them, most genes encode E3 ubiquitin ligases providing the substrate specificity to the pathway. However, due to this large number, identification on how ubiquitination mediates various effects in the cell is complicated. Instead, the study of ubiquitin receptors, " the interpreters of the ubiquitin signal", could bring further knowledge of the multiplicity of ubiquitin functions (Elsasser and Finley, 2005).

2.4.2 Mutation in the ubiquitin receptor DA1 leads to increased organ size

Many studies have shown that ubiquitin-proteasome pathway regulates plant development (for a review; (Brooks, 1973)). A dominant negative point mutation in an ubiquitin receptor, DA1 (large in Chinese), leads to the formation of larger organs in Arabidopsis (Li et al., 2008). Originally, this dominant negative DA1 (*da1-1*) allele was found in a genetic screen searching for mutants affecting organ size. *da1-1* produces larger seeds, embryos, ovules, petals, and siliques (Figure 2.7). *da1-1* leaves are also larger, rounder and contain more cells. The *da1-1* allele has a nonsynonymous substitution at a conserved amino acid at position 358 changing arginine to lysine. In contrast to *da1-1* mutant, a DA1 knock out line do not show *da1-1* phenotype implying that loss-of-function is not the cause of *da1-1* phenotype. Instead, overexpression of the point mutation of *da1-1*, $DA1^{R358K}$, shows similar phenotype as *da1-1* suggesting that this conserved single amino acid change causes the *da1-1* phenotype although the ubiquitin binding of $DA1^{R358K}$ is not affected. A kinematic analysis showed that the enlarged organ size is a result of a prolonged proliferation phase. *DA1* is highly expressed during the early stages of plant development to limit the period of cell proliferative, thereby it controls the time needed for most organs to reach their final size.

Figure 2.7 Organ size in *da1-1*. Nine-day-old seedlings of Col-0 (*A*) and *da1-1* (*B*). (*C*) The fifth leaves of Col-0 (*left*) and *da1-1* (*right*). Flowers of Col-0 (*D*) and *da1-1* (*E*). Siliques of Col-0 (*F*) and *da1-1* (*G*). Bars: *A*, *B*, *D–G*, 1 mm; *C*, 0.5 cm (from Li et al., 2008).

2.4.3 DA-related (DAR) proteins

Homologs of DA1, except for the ubiquitin-related motif region, have been discovered in crop plants such as rice and maize, but not in animals implying that this family of proteins has only evolved in plants (Li et al., 2008). The Arabidopsis genome encodes seven other proteins that share high similarity in amino acids with DA1 protein and have been called DA1-related (DAR) proteins (Li et al., 2008). DAR1 and DAR2 are the closest DA1 family members in Arabidopsis (Peng et al., 2015) and have been recently shown to control organ growth with *DA1* via modulation of endoreduplication (Peng et al., 2015). *DA1*, *DAR1*, and *DAR2* show similar expression patterns during leaf development in Arabidopsis suggesting overlapping functions in leaf growth. While *da1-ko1dar1-1* double mutants produce larger seeds and leaves, two other double mutants *da1-ko1dar 2-1* and *dar1-1dar2-1* do not exhibit any alteration in organ size (Li et al., 2008; Xia et al., 2013; Peng et al., 2015). Interestingly, the triple mutant *da1-ko1dar1-1dar2-1* form very small plants in comparison to wild type that can be complemented by overexpression of one of the three genes. Overexpression DA1, DAR1, or DAR2 in wild type does not affect organ size. Based on these observations, it was suggested that DA1, DAR1, and DAR2 act redundantly to control plant development in an organ-dependent and in a context-dependent manner. The triple mutant showed very small plants and leaves with reduced cell size which is associated with decreased ploidy levels. Through Yeast-two-hybrid screening, TEOSINTE BRANCHED1/CYCLOIDEA/PCF (TCP) 15 and TCP14, which is the closest homolog of TCP15, were identified as DA1 binding proteins. TCP15 and TCP14 are transcription factors and they are known to regulate endoreduplication (Kieffer et al., 2011; Li et al., 2012). DAR1 and DAR2 also interact with TCP15 and TCP14. These interactions modulate the stability of TCP14/15 and in the triple mutant *da1-ko1 dar 1-1 dar2-1,* protein levels of TCP14/15 were high. The pentuple *tcp14tcp15da1-ko1dar1-1dar2-1* mutant shows a higher endoreduplication index than *da1-ko1dar1-1dar2-1* triple mutant in the leaf suggesting that DA1, DAR1, and DAR2 act in a common pathway with TCP14/15 for the regulation of endoreduplication during leaf growth. The transcript levels of *RETINOBLASTOMA RELATED* (*RBR*) and *CYCA2;3*, known downstream targets of TCP14/15 and negative regulators of endoreduplication are dramatically increased in *da1-ko1 dar1-1dar2-1* triple mutant. This indicates that DA1, DAR1, and DAR2 are involved in the regulation of endoreduplication by modulating the stability of the transcription factors TCP14/15 that control the expression of key cell-cycle genes (Figure 2.8).

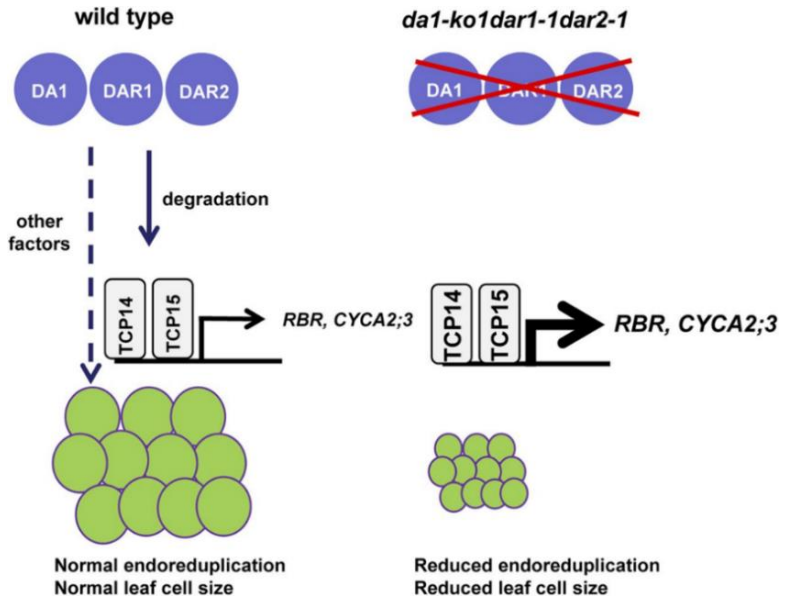


Figure 2.8 A model for DA1, DAR1, and DAR2 function in endoreduplication. DA1, DAR1, and DAR2 act redundantly to regulate endoreduplication by modulating the stability of the transcription factors TCP14/15. TCP14/15 repress endoreduplication by directly regulating the expression of several cell-cycle genes (e.g., *RBR* and *CYCA3;2*). Other factors might also mediate the effects of DA1, DAR1, and DAR2 on endoreduplication. In *da1-ko1 dar1-1 dar2-1* mutant, the transcription factors TCP14/15 are accumulating, resulting in increased expression of RBR and CYCA2;3 and strong repression of endoreduplication (from Peng et al., 2015).

2.4.4 ENHANCER OF DA1 (EOD1)/BIG BROTHER, an ubiquitin ligase

Via a genetic screen searching for modifiers of *da1-1*, one of the modifiers, enhancer of *da1-1* or EOD1-1 was identified (Li et al., 2008). *eod1-1 da1-1* double mutant exhibits a synergistic enhancement in organ size and has a longer life span than *da1-1* (Figure 2.10). It turned out that *EOD1* corresponds to the *BIG BROTHER* (*BB*), a gene encoding an E3 ubiquitin ligase previously described to inhibit organ growth (Disch et al., 2006). *eod1-1* allele also leads to a single amino acid change in BB protein. The original *bb-1* mutant and a T-DNA insertion line of EOD1 could enhance the phenotype of *da1-1* allele. It has been suggested that DA1 and EOD1/BB act in parallel pathways and have a common target to control organ size by modulating the duration of cell proliferation (Li et al., 2008) (Figure 2.9).

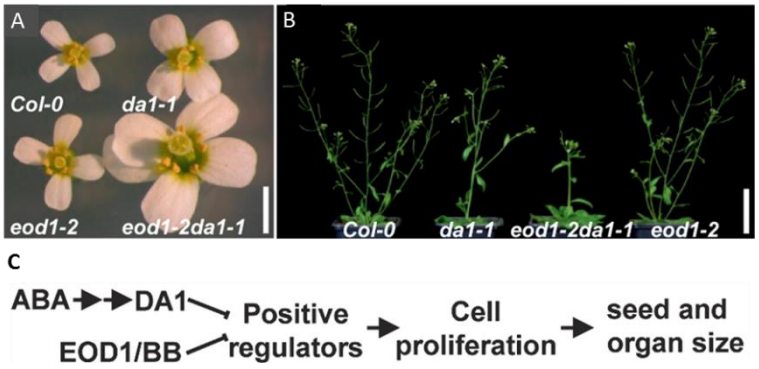


Figure 2.9 Mutations in *EOD1/BB* synergistically enhance the phenotypes of *da1-1*. Flowers (A) and plants (B) of Col-0, *da1-1*, *eod1-2da1-1* double mutant, and *eod1-2* are shown. (C) Model of *DA1* and *EOD1/BB* regulating seed and organ size through modulating cell proliferation (from Li et al., 2008).

2.4.5 DA2, an ubiquitin ligase

Through a screen of Arabidopsis T-DNA insertion lines of predicted ubiquitin ligases expressed in ovules and/or seeds, one mutant with altered seed size was identified and named *da2-1* (Xia et al., 2013). *DA2* encodes a REALLY INTERESTING NEW GENE (RING) type protein having E3 ubiquitin ligase activity. *da2-1* mutant produces larger and heavier seeds as well as bigger flowers, larger leaves, and has increased biomass compared to wild type. The larger leaf area is a result of increased cell number. Overexpressing plants of *DA2* show small leaves, reduced plant height, and decreased biomass as well as small flowers and short siliques supporting that DA2 negatively regulates organ size. Furthermore, *da2-1* mutant significantly enhances the phenotype of *da1-1* by further increasing seed size, cotyledon size and cell number. However, *da2-1 eod 1-2* double mutant displays additive effect suggesting that DA2 acts independently of *eod1-2*. The expression pattern of *DA2* is similar to *DA1* with high expression in younger organs such as leaf primordia and younger ovules and also in roots. *In vivo* and *in vitro* protein-protein interaction assays showed that DA1 interacts physically with DA2 through its C-terminal region. Homologues of DA2 were found in crops such as oilseed rape, soybean, rice, maize, and barley. In rice, this homologue is called GW2 and shares 43.1% identity with the amino acid sequence of the Arabidopsis DA2. Overexpression of GW2 in rice leads to reduced grain width (Song et al., 2007), formation of smaller organs and seeds implying a possible conserved function for these two proteins for organ growth regulation in Arabidopsis and rice (Xia et al., 2013).

2.4.6 UBP15/SOD2: ubiquitin-specific protease

One of the downstream target of DA1, *SUPPRESSOR2 OF DA1* (*SOD2*) was identified through an EMS screen searching for modifiers of the seed size phenotype of *da1-1* (Du et al., 2014). *sod2-1* mutation was mapped in a gene encoding UBIQUITIN-SPECIFIC PROTEASE 15 (UBP15) which has a role in protein de-ubiquitination and is known to regulate leaf development (Liu et al., 2008). The mutation in *sod2-1* allele is located at a site causing an alteration of splicing. *UBP15* is highly expressed in young organs. UBP15 physically interacts with DA1 and this interaction induces degradation of UBP15 suggesting that DA1 modulates the stability of UBP15 protein. *ubp15-1* mutant exhibits reduced seed and seedling size with reduced cell number. *ubp15-1*/*da1-1* double mutant has comparable phenotype as *ubp15-1* single mutant. In contrast, plants overexpressing *UBP15* produce large seeds and organs similar to *da1-1*. It has been therefore suggested that *ubp15-1* is epistatic to *da1-1* and *DA1* and *UBP15* play antagonistic roles in a common pathway controlling cell proliferation and thereby organ size (Figure 2.10). In addition, double mutants with other genes that enhance *da1-1* phenotype, *ubp15-1 eod1-2* and *ubp15-1 da2-1* show additive phenotype for both seed size and leaf area implying that UBP15 acts independently of *DA2* and *EOD1.*

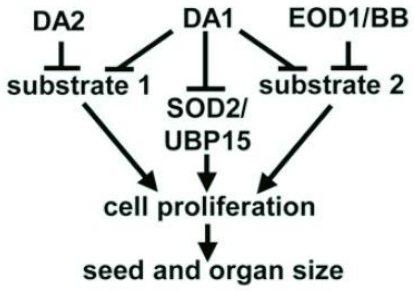


Figure 2.10 Genetic and molecular model of DA1, DA2, EOD1/BB, and UBP15/SOD2-mediated regulation of seed and organ size. DA1 and DA2 act synergistically to restrict seed and organ size, suggesting that DA1 and DA2 might share a common downstream substrate. Similarly, DA1 and EOD1 may have a common substrate. However, DA2 acts independently of EOD1 to affect seed and organ size, suggesting that DA2 and EOD1 could target different cell proliferation stimulators (indicated as substrate 1 and substrate 2) for degradation, with common regulation via DA1. DA1 acts upstream of UBP15 and controls its stability. However, UBP15 acts independently of both DA2 and EOD1, suggesting that UBP15 is not the target of DA2 or EOD1 for degradation (from Du et al., 2014).

2.4.7 EOD8/MEDIATOR COMPLEX SUBUNIT 25

Several general regulatory machineries have distinct roles in the control of organ size, the mediator complex being one of those. The mediator complex has a cellular housekeeping role by transferring diverse signals from transcription factors, activators, and repressors to the RNA polymerase II machinery to begin transcription (Kim et al., 1994; Koleske and Young, 1994). Interestingly, MEDIATOR COMPLEX SUBUNIT 25 (MED 25), a sub-unit of the mediator complex, was discovered as one of the enhancers of *da1-1* phenotype and the mutation was denominated *eod8-1* (Xu and Li, 2011). *eod8-1*/*med 25-1* has a point mutation at a highly conserved splicing site resulting in a translational frameshift and a premature translational product. *med25* mutants produce larger floral organs and younger leaves with an increased cell area and slightly increased cell number compared to wild type. In contrast, MED25 overexpression inhibits organ growth since narrower and shorter leaves, slightly smaller flowers, and thinner stems with decreased cell number and size are produced. Increased expression of *EXPANSIN* genes was observed in *med25* mutant suggesting that this could be a reason for the larger cells phenotype in *med25* mutant. *med25-2 da1-1* double mutant showed no changes in cell size but cell number of petals was increased. These analyses showed that MED25 functions redundantly with DA1 to control organ growth by restricting cell proliferation. A model for the action of MED25 has been suggested in which MED25 activates the transcription of target genes, after receiving growth signals. This MED25 mediated activation of gene expression results in a decreased cell expansion and cell proliferation, the latter acting in synergy with DA1 (Figure 2.11).

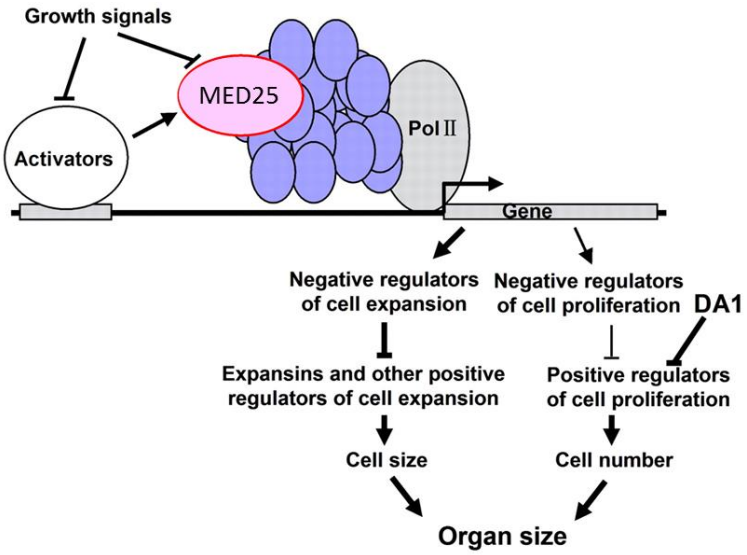


Figure 2.11 Model of *MED25* control of organ size. Growth signals are transmitted to the Mediator complex by direct action on MED25 or via activators to regulate the transcription of target genes, which include negative regulators of both cell proliferation and cell expansion. These negative regulators of cell expansion may regulate cell size by repressing the expression of *EXPANSIN* genes and other positive regulators of cell expansion. DA1 and MED25 might share a common downstream target to restrict cell proliferation (from Xu and Li, 2011).

References

Aach, H., Bode, H., Robinson, D.G., and Graebe, J.E. (1997). ent-Kaurene synthase is located in proplastids of meristematic shoot tissues. Planta **202:** 211-219.

Bai, Y., Meng, Y., Huang, D., Qi, Y., and Chen, M. (2011). Origin and evolutionary analysis of the plant-specific TIFY transcription factor family. Genomics **98:** 128-136.

Biemelt, S., Tschiersch, H., and Sonnewald, U. (2004). Impact of altered gibberellin metabolism on biomass accumulation, lignin biosynthesis, and photosynthesis in transgenic tobacco plants. Plant Physiol **135:** 254-265.

Brooks, K.E. (1973). Reproductive Biology of Selaginella: I. Determination of Megasporangia by 2-Chloroethylphosphonic Acid, an Ethylene-releasing Compound. Plant Physiol **51:** 718-722.

Carrera, E., Bou, J., Garcia-Martinez, J.L., and Prat, S. (2000). Changes in GA 20-oxidase gene expression strongly affect stem length, tuber induction and tuber yield of potato plants. Plant J **22:** 247-256.

Carvalhais, L.C., Dennis, P.G., Badri, D.V., Tyson, G.W., Vivanco, J.M., and Schenk, P.M. (2013). Activation of the jasmonic acid plant defence pathway alters the composition of rhizosphere bacterial communities. PLoS ONE **8:** e56457.

Century, K., Reuber, T.L., and Ratcliffe, O.J. (2008). Regulating the regulators: the future prospects for transcription-factor-based agricultural biotechnology products. Plant Physiol **147:** 20-29.

Chandler, P.M., Marion-Poll, A., Ellis, M., and Gubler, F. (2002). Mutants at the Slender1 locus of barley cv Himalaya. Molecular and physiological characterization. Plant Physiol **129:** 181-190.

Chang, C., and Bleecker, A.B. (2004). Ethylene biology. More than a gas. Plant Physiol **136:** 2895-2899.

Chini, A., Fonseca, S., Chico, J.M., Fernández-Calvo, P., and Solano, R. (2009). The ZIM domain mediates homo- and heteromeric interactions between Arabidopsis JAZ proteins. Plant J. **59:** 77-87.

Choi, D., Lee, Y., Cho, H.T., and Kende, H. (2003). Regulation of expansin gene expression affects growth and development in transgenic rice plants. Plant Cell **15:** 1386-1398.

Claeys, H., De Bodt, S., and Inzé, D. (2014). Gibberellins and DELLAs: central nodes in growth regulatory networks. Trends Plant Sci. **19:** 231-239.

Coles, J.P., Phillips, A.L., Croker, S.J., García-Lepe, R., Lewis, M.J., and Hedden, P. (1999). Modification of gibberellin production and plant development in *Arabidopsis* by sense and antisense expression of gibberellin 20-oxidase genes. Plant J. **17:** 547-556.

Cuellar Perez, A., Nagels Durand, A., Vanden Bossche, R., De Clercq, R., Persiau, G., Van Wees, S.C., Pieterse, C.M., Gevaert, K., De Jaeger, G., Goossens, A., and Pauwels, L. (2014). The non-JAZ TIFY protein TIFY8 from Arabidopsis thaliana is a transcriptional repressor. PLoS ONE **9:** e84891.

Dai, C., and Xue, H.W. (2010). Rice early flowering1, a CKI, phosphorylates DELLA protein SLR1 to negatively regulate gibberellin signalling. EMBO J **29:** 1916-1927.

Davies, P. (2012). Plant hormones and their role in plant growth and development. (Springer Science & Business Media).

de Lucas, M., Davière, J.-M., Rodríguez-Falcón, M., Pontin, M., Iglesias-Pedraz, J.M., Lorrain, S., Fankhauser, C., Blázquez, M.A., Titarenko, E., and Prat, S. (2008). A molecular framework for light and gibberellin control of cell elongation. Nature **451:** 480-484.

Dewitte, W., Scofield, S., Alcasabas, A.A., Maughan, S.C., Menges, M., Braun, N., Collins, C., Nieuwland, J., Prinsen, E., Sundaresan, V., and Murray, J.A.H. (2007). *Arabidopsis* CYCD3 D-type cyclins link cell proliferation and endocycles and are rate-limiting for cytokinin responses. Proc. Natl. Acad. Sci. USA **104:** 14537-14542.

Dill, A., Jung, H.-S., and Sun, T.-p. (2001). The DELLA motif is essential for gibberellin-induced degradation of RGA. Proc. Natl. Acad. Sci. USA **98:** 14162-14167.

Dill, A., Thomas, S.G., Hu, J., Steber, C.M., and Sun, T.-p. (2004). The Arabidopsis F-box protein SLEEPY1 targets gibberellin signaling repressors for gibberellin-induced degradation. Plant Cell **16:** 1392-1405.

Disch, S., Anastasiou, E., Sharma, V.K., Laux, T., Fletcher, J.C., and Lenhard, M. (2006). The E3 ubiquitin ligase BIG BROTHER controls *Arabidopsis* organ size in a dosage-dependent manner. Curr. Biol. **16:** 272-279.

Du, L., Li, N., Chen, L., Xu, Y., Li, Y., Zhang, Y., Li, C., and Li, Y. (2014). The ubiquitin receptor DA1 regulates seed and organ size by modulating the stability of the ubiquitin-specific protease UBP15/SOD2 in *Arabidopsis*. Plant Cell **26:** 665-677.

Elsasser, S., and Finley, D. (2005). Delivery of ubiquitinated substrates to protein-unfolding machines. Nat Cell Biol **7:** 742-749.

Elsner, J., Michalski, M., and Kwiatkowska, D. (2012). Spatiotemporal variation of leaf epidermal cell growth: a quantitative analysis of *Arabidopsis thaliana* wild-type and triple *cyclinD3* mutant plants. Ann. Bot. **109:** 897-910.

Eriksson, M.E., Israelsson, M., Olsson, O., and Moritz, T. (2000). Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. Nat. Biotechnol. **18:** 784-788.

Fariduddin, Q., Yusuf, M., Ahmad, I., and Ahmad, A. (2014). Brassinosteroids and their role in response of plants to abiotic stresses. Biol. Plant. **58:** 9-17.

Feng, S., et al. (2008). Coordinated regulation of *Arabidopsis thaliana* development by light and gibberellins. Nature **451:** 475-479.

Finkelstein, R. (2013). Abscisic acid synthesis and response. The Arabidopsis Book**:** e0166.

Fu, X., Richards, D.E., Fleck, B., Xie, D., Burton, N., and Harberd, N.P. (2004). The Arabidopsis mutant sleepy1$^{gar2\text{-}1}$ protein promotes plant growth by increasing the affinity of the SCF^{SLY1} E3 ubiquitin ligase for DELLA protein substrates. Plant Cell **16:** 1406-1418.

Gonzalez, N., Beemster, G.T.S., and Inzé, D. (2009). David and Goliath: what can the tiny weed *Arabidopsis* teach us to improve biomass production in crops? Curr. Opin. Plant Biol. **12:** 157-164.

Gonzalez, N., et al. (2010). Increased leaf size: different means to an end. Plant Physiol. **153:** 1261-1279.

Gonzalez, N., et al. (2015). A Repressor Protein Complex Regulates Leaf Growth in Arabidopsis. Plant Cell **27:** 2273-2287.

Guo, M., et al. (2014). Maize ARGOS1 (ZAR1) transgenic alleles increase hybrid maize yield. J Exp Bot **65:** 249-260.

Ha, S., Vankova, R., Yamaguchi-Shinozaki, K., Shinozaki, K., and Tran, L.-S.P. (2012). Cytokinins: metabolism and function in plant adaptation to environmental stresses. Trends Plant Sci. **17:** 172-179.

Harberd, N.P., Belfield, E., and Yasumura, Y. (2009). The angiosperm gibberellin-GID1-DELLA growth regulatory mechanism: how an "inhibitor of an inhibitor" enables flexible response to fluctuating environments. Plant Cell **21:** 1328-1339.

Hedden, P. (2003). The genes of the Green Revolution. Trends Genet. **19:** 5-9.

Hedden, P., and Phillips, A.L. (2000). Manipulation of hormone biosynthetic genes in transgenic plants. Curr Opin Biotechnol **11:** 130-137.

Hedden, P., and Thomas, S.G. (2012). Gibberellin biosynthesis and its regulation. Biochem. J. **444:** 11-25.

Helliwell, C.A., Sullivan, J.A., Mould, R.M., Gray, J.C., Peacock, W.J., and Dennis, E.S. (2001). A plastid envelope location of Arabidopsis ent-kaurene oxidase links the plastid and endoplasmic reticulum steps of the gibberellin biosynthesis pathway. Plant J **28:** 201-208.

Heo, J.-O., Chang, K.S., Kim, I.A., Lee, M.-H., Lee, S.A., Song, S.-K., Lee, M.M., and Lim, J. (2011). Funneling of gibberellin signaling by the GRAS transcription regulator SCARECROW-LIKE 3 in the *Arabidopsis* root. Proc. Natl. Acad. Sci. USA **108:** 2166-2171.

Hu, Y., Xie, Q., and Chua, N.-H. (2003). The Arabidopsis auxin-inducible gene ARGOS controls lateral organ size. The Plant Cell **15:** 1951-1961.

Hu, Y., Poh, H.M., and Chua, N.-H. (2006). The Arabidopsis *ARGOS-LIKE* gene regulates cell expansion during organ growth. Plant J. **47:** 1-9.

Huang, S., Raman, A.S., Ream, J.E., Fujiwara, H., Cerny, R.E., and Brown, S.M. (1998). Overexpression of 20-oxidase confers a gibberellin-overproduction phenotype in Arabidopsis. Plant Physiol. **118:** 773-781.

Ikeda, A., Ueguchi-Tanaka, M., Sonoda, Y., Kitano, H., Koshioka, M., Futsuhara, Y., Matsuoka, M., and Yamaguchi, J. (2001). slender rice, a constitutive gibberellin response mutant, is caused by a null mutation of the SLR1 gene, an ortholog of the height-regulating gene GAI/RGA/RHT/D8. Plant Cell **13:** 999-1010.

Israelsson, M., Sundberg, B., and Moritz, T. (2005). Tissue-specific localization of gibberellins and expression of gibberellin-biosynthetic and signaling genes in wood-forming tissues in aspen. The Plant Journal **44:** 494-504.

Kieffer, M., Master, V., Waites, R., and Davies, B. (2011). TCP14 and TCP15 affect internode length and leaf shape in Arabidopsis. Plant J. **68:** 147-158.

Kim, Y.J., Bjorklund, S., Li, Y., Sayre, M.H., and Kornberg, R.D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell **77:** 599-608.

King, K.E., Moritz, T., and Harberd, N.P. (2001). Gibberellins are not required for normal stem growth in Arabidopsis thaliana in the absence of GAI and RGA. Genetics **159:** 767-776.

Ko, J.-H., Yang, S.H., and Han, K.-H. (2006). Upregulation of an Arabidopsis RING-H2 gene, *XERICO*, confers drought tolerance through increased abscisic acid biosynthesis. Plant J. **47:** 343-355.

Koleske, A.J., and Young, R.A. (1994). An RNA polymerase II holoenzyme responsive to activators. Nature **368:** 466-469.

Lau, O.S., Davies, K.A., Chang, J., Adrian, J., Rowe, M.H., Ballenger, C.E., and Bergmann, D.C. (2014). Direct roles of SPEECHLESS in the specification of stomatal self-renewing cells. Science **345:** 1605-1609.

Lee, S., Cheng, H., King, K.E., Wang, W., He, Y., Hussain, A., Lo, J., Harberd, N.P., and Peng, J. (2002). Gibberellin regulates Arabidopsis seed germination via RGL2, a GAI/RGA-like gene whose expression is up-regulated following imbibition. Genes Dev **16:** 646-658.

Li, Y., Zheng, L., Corke, F., Smith, C., and Bevan, M.W. (2008). Control of final seed and organ size by the *DA1* gene family in *Arabidopsis thaliana*. Genes Dev. **22:** 1331-1336.

Li, Z.Y., Li, B., and Dong, A.W. (2012). The Arabidopsis transcription factor AtTCP15 regulates endoreduplication by modulating expression of key cell-cycle genes. Mol Plant **5:** 270-280.

Liu, Y., Wang, F., Zhang, H., He, H., Ma, L., and Deng, X.W. (2008). Functional characterization of the Arabidopsis ubiquitin-specific protease gene family reveals specific role and redundancy of individual members in development. Plant J **55:** 844-856.

McGinnis, K.M., Thomas, S.G., Soule, J.D., Strader, L.C., Zale, J.M., Sun, T.P., and Steber, C.M. (2003). The Arabidopsis SLEEPY1 gene encodes a putative F-box subunit of an SCF E3 ubiquitin ligase. Plant Cell **15:** 1120-1130.

Monna, L., Kitazawa, N., Yoshino, R., Suzuki, J., Masuda, H., Maehara, Y., Tanji, M., Sato, M., Nasu, S., and Minobe, Y. (2002). Positional cloning of rice semidwarfing gene, sd-1: rice "green revolution gene" encodes a mutant enzyme involved in gibberellin synthesis. DNA Res **9:** 11-17.

Nakajima, M., et al. (2006). Identification and characterization of Arabidopsis gibberellin receptors. Plant J **46:** 880-889.

Nelissen, H., Rymen, B., Jikumaru, Y., Demuynck, K., Van Lijsebettens, M., Kamiya, Y., Inzé, D., and Beemster, G.T.S. (2012). A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. Curr. Biol. **22:** 1183-1187.

Oikawa, T., Koshioka, M., Kojima, K., Yoshida, H., and Kawata, M. (2004). A role of OsGA20ox1 , encoding an isoform of gibberellin 20-oxidase, for regulation of plant stature in rice. Plant Mol Biol **55:** 687-700.

Olszewski, N., Sun, T.P., and Gubler, F. (2002). Gibberellin signaling: biosynthesis, catabolism, and response pathways. Plant Cell **14 Suppl:** S61-80.

Pauwels, L., et al. (2010). NINJA connects the co-repressor TOPLESS to jasmonate signalling. Nature **464:** 788-791.

Peng, J., Carol, P., Richards, D.E., King, K.E., Cowling, R.J., Murphy, G.P., and Harberd, N.P. (1997). The *Arabidopsis GAI* gene defines a signaling pathway that negatively regulates gibberellin responses. Genes Dev. **11:** 3194-3205.

Peng, J., et al. (1999). `Green revolution' genes encode mutant gibberellin response modulators. Nature **400:** 256-261.

Peng, Y., Chen, L., Lu, Y., Wu, Y., Dumenil, J., Zhu, Z., Bevan, M.W., and Li, Y. (2015). The ubiquitin receptors DA1, DAR1, and DAR2 redundantly regulate endoreduplication by modulating the stability of TCP14/15 in Arabidopsis. Plant Cell **27:** 649-662.

Phillips, A.L. (2010). Genetic and transgenic approaches to improving crop performance. In Plant Hormones (Springer), pp. 618-645.

Plackett, A.R., Powers, S.J., Fernandez-Garcia, N., Urbanova, T., Takebayashi, Y., Seo, M., Jikumaru, Y., Benlloch, R., Nilsson, O., and Ruiz-Rivero, O. (2012). Analysis of the developmental roles of the Arabidopsis gibberellin 20-oxidases demonstrates that GA20ox1,-2, and-3 are the dominant paralogs. The Plant Cell **24:** 941-960.

Radhakrishnan, I., Pérez-Alvarado, G.C., Parker, D., Dyson, H.J., Montminy, M.R., and Wright, P.E. (1997). Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. Cell **91:** 741-752.

Rivas-San Vicente, M., and Plasencia, J. (2011). Salicylic acid beyond defence: its role in plant growth and development. J Exp Bot **62:** 3321-3338.

Rojas, C.A., Hemerly, A.S., and Ferreira, P.C. (2010). Genetically modified crops for biomass increase. Genes and strategies. GM crops **1:** 137-142.

Santner, A., and Estelle, M. (2009). Recent advances and emerging trends in plant hormone signalling. Nature **459:** 1071-1078.

Sasaki, A., Itoh, H., Gomi, K., Ueguchi-Tanaka, M., Ishiyama, K., Kobayashi, M., Jeong, D.H., An, G., Kitano, H., Ashikari, M., and Matsuoka, M. (2003). Accumulation of phosphorylated repressor for gibberellin signaling in an F-box mutant. Science **299:** 1896-1898.

Sasaki, A., Ashikari, M., Ueguchi-Tanaka, M., Itoh, H., Nishimura, A., Swapan, D., Ishiyama, K., Saito, T., Kobayashi, M., Khush, G.S., Kitano, H., and Matsuoka, M. (2002). Green revolution: a mutant gibberellin-synthesis gene in rice. Nature **416:** 701-702.

Schwechheimer, C., and Willige, B.C. (2009). Shedding light on gibberellic acid signalling. Curr. Opin. Plant Biol. **12:** 57-62.

Silverstone, A.L., Ciampaglio, C.N., and Sun, T. (1998). The Arabidopsis RGA gene encodes a transcriptional regulator repressing the gibberellin signal transduction pathway. Plant Cell **10:** 155-169.

Silverstone, A.L., Tseng, T.-S., Swain, S.M., Dill, A., Jeong, S.Y., Olszewski, N.E., and Sun, T.-p. (2007). Functional analysis of SPINDLY in gibberellin signaling in Arabidopsis. Plant Physiol. **143:** 987-1000.

Smalle, J., and Vierstra, R.D. (2004). The ubiquitin 26S proteasome proteolytic pathway. Annu. Rev. Plant Biol. **55:** 555-590.

Song, W.C., Qian, Y., Sun, X., and Negishi, M. (1997). Cellular localization and regulation of expression of testicular estrogen sulfotransferase. Endocrinology **138:** 5006-5012.

Song, X.-J., Huang, W., Shi, M., Zhu, M.-Z., and Lin, H.-X. (2007). A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nat. Genet. **39:** 623-630.

Spielmeyer, W., Ellis, M.H., and Chandler, P.M. (2002). Semidwarf (sd-1), "green revolution" rice, contains a defective gibberellin 20-oxidase gene. Proc Natl Acad Sci U S A **99:** 9043-9048.

Sun, T.-p. (2011). The molecular mechanism and evolution of the GA-GID1-DELLA signaling module in plants. Curr. Biol. **21:** R338-R345.

Sun, T.P., and Kamiya, Y. (1994). The Arabidopsis GA1 locus encodes the cyclase ent-kaurene synthetase A of gibberellin biosynthesis. Plant Cell **6:** 1509-1518.

Thakur, J.K., Agarwal, P., Parida, S., Bajaj, D., and Pasrija, R. (2013). Sequence and expression analyses of KIX domain proteins suggest their importance in seed development and determination of seed size in rice, and genome stability in Arabidopsis. Mol. Genet. Genomics **288:** 329-346.

Tyler, L., Thomas, S.G., Hu, J., Dill, A., Alonso, J.M., Ecker, J.R., and Sun, T.P. (2004). Della proteins and gibberellin-regulated seed germination and floral development in Arabidopsis. Plant Physiol **135:** 1008-1019.

Ueguchi-Tanaka, M., Nakajima, M., Motoyuki, A., and Matsuoka, M. (2007). Gibberellin receptor and its role in gibberellin signaling in plants. Annu. Rev. Plant Biol. **58:** 183-198.

Ueguchi-Tanaka, M., Ashikari, M., Nakajima, M., Itoh, H., Katoh, E., Kobayashi, M., Chow, T.-y., Hsing, Y.-i.C., Kitano, H., Yamaguchi, I., and Matsuoka, M. (2005). *GIBBERELLIN INSENSITIVE DWARF1* encodes a soluble receptor for gibberellin. Nature **437:** 693-698.

Varbanova, M., et al. (2007). Methylation of gibberellins by Arabidopsis GAMT1 and GAMT2. Plant Cell **19:** 32-45.

Vidal, A.M., Gisbert, C., Talon, M., Primo-Millo, E., Lopez-Diaz, I., and Garcia-Martinez, J.L. (2001). The ectopic overexpression of a citrus gibberellin 20-oxidase enhances the non-13-hydroxylation pathway of gibberellin biosynthesis and induces an extremely elongated phenotype in tobacco. Physiol Plant **112:** 251-260.

White, D.W.R. (2006). *PEAPOD* regulates lamina size and curvature in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **103:** 13238-13243.

Willige, B.C., Ghosh, S., Nill, C., Zourelidou, M., Dohmann, E.M.N., Maier, A., and Schwechheimer, C. (2007). The DELLA domain of GA INSENSITIVE mediates the interaction with the GA INSENSITIVE DWARF1A gibberellin receptor of *Arabidopsis*. Plant Cell **19:** 1209-1220.

Xia, T., Li, N., Dumenil, J., Li, J., Kamenski, A., Bevan, M.W., Gao, F., and Li, Y. (2013). The ubiquitin receptor DA1 interacts with the E3 ubiquitin ligase DA2 to regulate seed and organ size in *Arabidopsis*. Plant Cell **25:** 3347-3359.

Xu, R., and Li, Y. (2011). Control of final organ size by Mediator complex subunit 25 in *Arabidopsis thaliana*. Development **138:** 4545-4554.

Yamaguchi, S. (2008). Gibberellin metabolism and its regulation. Annu. Rev. Plant Biol. **59:** 225-251.

Yang, F., et al. (2006). An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. Nature **442:** 700-704.

Zentella, R., Zhang, Z.-L., Park, M., Thomas, S.G., Endo, A., Murase, K., Fleet, C.M., Jikumaru, Y., Nambara, E., Kamiya, Y., and Sun, T.-p. (2007). Global analysis of DELLA direct targets in early gibberellin signaling in *Arabidopsis*. Plant Cell **19:** 3037-3057.

Zhang, Z.-L., Ogawa, M., Fleet, C.M., Zentella, R., Hu, J., Heo, J.-O., Lim, J., Kamiya, Y., Yamaguchi, S., and Sun, T.-p. (2011). SCARECROW-LIKE 3 promotes gibberellin signaling by antagonizing master growth repressor DELLA in Arabidopsis. Proc. Natl. Acad. Sci. USA **108:** 2160-2165.

Zhao, Y. (2010). Auxin biosynthesis and its role in plant development. Annu. Rev. Plant Biol. **61:** 49.

Chapter 3

Natural variation in plants

Youn-Jeong Nam[1,2], Dirk Inzé[1,2*] & Nathalie Gonzalez[1,2*]

[1]Department of Plant Systems Biology, VIB B-9052 Gent, Belgium, [2]Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Gent, Belgium
* These authors contributed equally to this work.

Contributions: Y.J.N. was the main author of this chapter. D.I. and N.G. contributed to the writing of this chapter.

Chapter 3. Natural variation in plants

In plants, natural populations within the same species embrace an enormous diversity in morphology, physiology, response to environmental changes or disease susceptibility (Alonso-Blanco et al., 2009). Throughout history, humans have used this natural variation in more than 100 plant species for improvement and domestication (Diamond, 2002) showing that natural variation is a useful source for valuable traits for plant breeding. This developmental diversity results from naturally occurring genetic variations which have led to plant evolution through natural selection (Cronk, 2001). Quantitative traits, such as leaf growth, are generally driven by molecular polymorphisms at multiple loci and genes, which are called quantitative trait loci (QTL) and quantitative trait genes (QTGs), respectively (Alonso-Blanco et al., 2009). By studying natural variants, numerous genes and functional polymorphisms have been identified for different traits such as flowering time in many species including Arabidopsis, tomato, pea, barley, rice, and maize (Weigel, 2012; Huang and Han, 2014).

3.1 Studies using natural variation in Arabidopsis

Arabidopsis thaliana is a well suited model plant for studying natural variation since it has high homozygosity in natural populations because of a very high selfing rate and also a world-wide distribution in a wide range of ecological environments (Figure 3.1). These different ecological habitats made that in the past, Arabidopsis natural populations were referred as "ecotype" (Turesson, 1922). Nowadays, a neutral term, "accession", is used and refers to germplasm collections (Alonso-Blanco and Koornneef, 2000). Arabidopsis accessions show a remarkable range of phenotypic variations such as for vegetative rosette size as illustrated in Figure 3.1. Thus far, natural variation has been studied for a number of phenotypic traits and a number of causative genes for the phenotype investigated have been cloned in Arabidopsis (Weigel, 2012).

Figure 3.1 Global geographic distribution of *Arabidopsis thaliana* in the wild. Areas colored in red represent the geographic distribution of Arabidopsis. Black dots show the origin of 17 accessions as examples to highlight the variability in rosette phenotype. The pictures represent 25 day-old rosette of these 17 accessions (modified from Kramer, 2015).

The study of natural variation requires the integration of two types of information: genotypic data (e.g. genome sequences or molecular markers) and phenotypic data (traits measurements). In plants, genetic mapping or QTL mapping, is the most common method for using natural variation and needs at least two or more strains containing different alleles (Alonso-Blanco and Koornneef, 2000). By crossing these strains, an artificial population is generated such as recombinant inbred lines (RILs; population of plants containing recombinant chromosomes from two parental lines (Reiter et al., 1992)) or multiple advanced generation intercross (MAGIC) population (Kover et al., 2009) and Arabidopsis multi-parent RIL (AMPRIL) ((Huang et al., 2011); populations made by a combination of multiple parents). After generating multiple successive generations, the phenotypes of the derived population are scored. Specific traits are then associated with segregating genetic markers. Markers that are genetically linked to a QTL influencing the trait of interest will show significant association with a specific phenotype while unlinked genetic markers will barely segregate with the trait values. Usually, the more generations of recombinant inbred strains are generated, the greater is the power and resolution with which phenotypes can be mapped to chromosomal locations since the size of recombination fragments is more reduced. Although multi parental lines provide more diversity in genotype, it still has a limitation of accuracy because of the presence of many linked genes that are not separated by recombination.

Genome-wide association studies (GWAS) can be an excellent complement to QTL mapping. GWAS search for on an association between unlinked individual genes or nucleotides such as single-nucleotide polymorphisms (SNPs) and phenotypic data rendering GWAS more precise for finding a gene linked to phenotype than QTL mapping. Furthermore, the population used in GWAS corresponds to unrelated individuals such as Arabidopsis accessions. However, GWAS also has limitations since it can generate a large number of false positives and this study only can be applied on an organism which has enough genomic resources. By conducting both QTL mapping and GWAS, specific genes that are responsible for a phenotype of interest can potentially be identified (Miles and Wayne, 2008).

3.2 Arabidopsis natural variation to study growth

Plant growth is an important trait for agriculture since it is directly related to crop yield. Thus, it is valuable to dissect this multi-factorial trait and understand its regulation. Growth can be described in many ways depending on which parameters are measured (weight, area, length) and at which level (plant, leaf or cellular level) (Dhondt et al., 2013). Furthermore, since growth is a dynamic process, time-component parameters need to be assessed such as relative growth rate (RGR) providing the speed of plant growth (Dhondt et al., 2014).

Arabidopsis natural variation for growth and growth-related parameters has been described in several studies. For example, studies have focused on leaf architecture (Pérez-Pérez et al., 2002), plant height (Pigliucci and Schlichting, 1995), fresh weight or dry weight (Aarssen

and Clauss, 1992; Li et al., 1998; Loudet et al., 2003; El-Lithy et al., 2004; Meyer et al., 2007; Sulpice et al., 2013), relative growth rate of the rosette (Li et al., 1998; El-Lithy et al., 2004), total leaf number (Li et al., 1998; El-Lithy et al., 2004; Massonnet et al., 2011), rosette area (Li et al., 1998; El-Lithy et al., 2004; Massonnet et al., 2011; Clauw et al., 2015), cell expansion rate (Beemster et al., 2002; Massonnet et al., 2011), cell number (Massonnet et al., 2011; Clauw et al., 2015), cell area (Massonnet et al., 2011; Clauw et al., 2015), endoreduplication index (EI) (Massonnet et al., 2011), stomatal index (Clauw et al., 2015) or hypocotyl length (Millenaar et al., 2005).

In these studies, the range of variability of parameter values differed depending on which parameters are measured. For example, among 22 natural accessions, rosette area showed 75% of variation between the smaller and the larger plants, while dry weight of plants had a 96% variation (El-Lithy et al., 2004). Importantly, the range of values can be influenced by two factors: the origin of the genotypes used and the number of accessions analysed. Forty accessions which originated from only European countries showed 43% of variation in dry weight (Li et al., 1998). However, using 22 accessions selected from 130 accessions based on obvious differences in growth characteristics such as biomass (both fresh and dry weight) of above ground plant and seed weight and wide range of origins, 96% of variation in dry weight was observed (El-Lithy et al., 2004).

Measurements of growth-related parameters in natural variants can be used to calculate correlations in order to gain insight into potential relations between parameters or potential contribution of a parameter to a higher phenotype level, such as how a cellular parameter (cell number or cell size) can influence organ level output (rosette biomass or leaf size). For example, a correlation analysis between biomass and specific metabolic components suggested that biomass can be predicted by a set of low-molecular weight metabolites (Meyer et al., 2007; Sulpice et al., 2009; Sulpice et al., 2010; Sulpice et al., 2013) such as starch which is negatively correlated with biomass in 94 accessions (Sulpice et al., 2009). A positive correlation was found between dry weight and relative growth rate (RGR) of the leaf among 40 accessions (Li et al., 1998). However, correlation analysis between growth-related parameters can be affected by the population used for the measurements. For example, a positive correlation was found between endoreduplication levels and cell number in an endoreduplication mutant collection whereas endoreduplication positively correlated with cell area when using a RILs population between Ler-0 and An-1 accessions (Massonnet et al., 2011).

Natural variation can be used for phenotypic screening as an alternative to classical genetic approach to identify genetic defects such as a mutation in a specific accession. This natural variation screening method was for example used to identify genetic components important for environmental perturbation response. By growing accessions at 27°C and in a short-day condition, leaf growth abnormalities that were referred to as "*irregularly impaired leaves*" (*iil*) were observed in Bur-0. Via linkage mapping (similar to QTL mapping but working with a single gene traits or disease) and using a RIL population between Bur-0 and Pf-0 (an

accession that does not show the abnormal leaf phenotype), the locus of *IIL* was mapped. *IIL* encodes one of the leucine biosynthesis enzymes and the expression of *IIL* is interrupted in Bur-0 due to variation in the length of simple DNA triplet repeats, therefore causing the specific leaf phenotype at high temperature (Sureshkumar et al., 2009).

To examine the light-signalling pathway, 141 accessions were screened by measuring hypocotyl length in response to four different light conditions and two hormones, GA and BR, known to interact with light-signalling pathway (Maloof et al., 2001). Some accessions showed similar responses as known light-signalling mutants while others displayed new patterns of responses. Via western blot, the expression and activity of photoreceptors such as PHYTOCHROME A which is degraded by light were analysed. It was shown that the decreased sensitivity to light in Lm-2 accession results from a mutation in *PHYTOCHROME A* leading to higher accumulation of this protein.

In summary, natural variation has been used widely not only through GWAS and QTL mapping but also for phenotypic screening caused by naturally occurring mutations to define novel molecular functions. Natural variation therefore provides a good genetic resource to study and unravel the mechanism underlying plant development.

3.3 Natural variation and gene expression

RNA transcript is the first intermediate between the genome and the final phenotype. Thus, transcriptional differences could be an important factor that causes phenotypic variation in natural variants. To understand how transcript variation affects downstream phenotypic trait variation, transcript profiling at the genome scale is required. By considering transcript levels as quantitative traits, variation in these levels can be used to map expression QTL (eQTL) (Schadt et al., 2003) or to perform expression GWAS (eGWAS) (Keurentjes et al., 2007). The genetic changes leading to phenotypic variation (transcripts in this case) can be found in two different regions in the genome. If genetic variation occurs in a coding region, it might change the function of the encoded protein. Genetic variation could affect the expression pattern of a certain gene when located in a *cis* regulatory (e.g promoter sequence) or in a *trans* acting region (e.g. a transcription factor). Through a large global eQTL mapping in a RIL (Sha x Bay-0) population of 211 individuals, a highly variable and complex transcript-level variation was observed in Arabidopsis (West et al., 2007). Almost 70 % of the transcripts (e-traits) were mapped in single eQTL. More than 30 % of transcripts were regulated by *cis*-eQTL and numerous *trans*-eQTL were regulating hundreds to thousands of transcripts suggesting a complex genetic architecture of transcript-level variation in Arabidopsis.

Natural variation in response to environmental perturbation such as plant hormone treatment or mild drought has been investigated at transcriptional level to identify molecular mechanisms explaining phenotypic variation. Seven accessions were treated with salicylic acid (SA), a key signalling molecule involved in plant-pathogen interactions, and transcriptome

response was analyzed (van Leeuwen et al., 2007). Differentially expressed genes were searched in the different accessions to detect natural variation in gene expression associated with SA response at individual genes and gene networks level. Almost 4000 genes displayed significant differential expression and interestingly, 95 % of those genes were showing accession-specific response. A significant variation at gene-network level was also detected between the seven accessions implying that genetic variation exists upstream of SA signal transduction.

Physiological and global transcriptome variation in response to exogenous auxin treatment, one of the key regulators of plant development, was also measured in seven accessions (Delker et al., 2010). Highly accession-specific transcriptional variations were detected between the accessions. However, this variation was not caused by sequence diversity in the auxin signalling genes whose gene products, the TIR1/AFB auxin receptors, the Aux/IAA auxin signalling repressors, and the ARF transcription factors, require direct physical interactions between them for regulation of auxin responses. Instead, sequence polymorphism in *cis*- or *trans*-acting factors affecting the regulation of auxin signalling gene expression might cause differences in the amount of interacting signalling components. This quantitative distortions of auxin signalling components leads to either repressing or promoting of auxin signalling resulting in the natural variation of the downstream transcriptional responses.

The analysis of genome-wide transcriptome response to mild drought stress in young developing leaves identified 354 genes with common differential expression pattern in six accessions (Clauw et al., 2015). Only few genes were found to have an accession-specific response in young leaves and none of those genes were differentially expressed in only one accession. This result demonstrated that a robust set of genes function in the mild drought response in different genetic backgrounds in early developmental stages.

In summary, intraspecific transcriptome variation upon developmental and environmental stimuli in Arabidopsis is complex. The response can be either highly accession-specific or a robust response over different genetic backgrounds depending on the perturbation applied can be found.

3.4 Transgenic approaches in natural variants

Arabidopsis is easily transformable with the floral dip method (Clough and Bent, 1998) and transgenic approaches -overexpression or silencing- has been used in different accessions mainly to validate the identification by QTL or GWAS of a putative causative gene for a specific trait via complementation (Wang et al., 2007; Boggs et al., 2009; Huang et al., 2012; Meijón et al., 2014; Miller et al., 2015). In addition, transgenic approaches have been used to monitor the expression patterns of target genes with a marker line in various accessions. In order to examine how daily rhythms fluctuate in natural variants and how genetic variation influences their phenotypes, the promoter of the evening-expressed gene *GIGANTEA* (*GI*), which reflects daily rhythms, fused to the luciferase marker gene was introduced in 77 accessions (de Montaigu et

al., 2015). The influence of natural variation in the timing of *GI* expression was observed mostly in long day condition. QTL analysis using Col-0 and Lip-0 showing opposite expression patterns of *GI* revealed *TIMING OF GI* (*TOG*) alleles that regulate the timing of *GI* expression. *PHYTOCHROME B (PHYB)* gene, encoding the red light photoreceptor, was identified as one of the *TOG* alleles. It has been shown that GI inhibits hypocotyl growth activated by PHYB and represses the expression of *PHYTOCHROME INTERACTING FACTOR* 4 (*PIF4*), encoding a transcription factor. A synergistic action was observed between GI and PHYB to inhibit hypocotyl growth and repress *PIF4* in *phyB-9/gi-2* double mutant. A strong positive correlation was found between hypocotyl length and *PIF4* mRNA levels. Therefore, the modified expression of *GI* by *TOG* alleles at least partially activated by PHYB activity could sufficiently modulate the expression of *PIF4* and lead to physiological growth changes in the different accessions.

Thus, transgenic approaches in natural variants have been used as a tool not only for complementation study but also for checking the dynamics of gene expression of a particular gene.

3.5 Genetic background effects, some general concepts

When transgenes or mutations are introduced into different accessions, genetic background effect should be considered since various mutational variations might influence the effect of the gene introduced (Chandler et al., 2013; Paaby and Rockman, 2014). This phenomenon can also be described as epistasis since the effect of a gene is dependent on interactions with one or more modifier genes, in this case, the gene variants from the different genetic backgrounds (Figure 3.2A).

Different genetic backgrounds might either restrict or pronounce the expression of the same mutation because of interaction with different genes therefore leading to variable phenotypes. These variable phenotypes which refer to as penetrance can be measured by calculating the proportion of individuals that actually show the expected phenotypes from the population. In this thesis, I use “expressivity” as a term that describes the degree of leaf size alteration in a transgenic line compared to its corresponding wild type. For example, if the expressivity of a certain genetic perturbation is high in an accession, it means that the transgenic lines of this accession show increased leaf area. In contrast, low expressivity means decreased leaf area compared to wild type.

There are many steps between a specific genotype and the output at phenotypic level where another gene product can interfere and mediate differences in penetrance and expressivity (Figure 3.2B). Transcriptional, translational or post-translational regulators can influence a final phenotype by affecting transcription rate, protein synthesis or conformation leading to more or less activity of the protein. Therefore, between accessions, any differences in gene sequences that can influence the regulation of transcription and/or activity of a protein can alter its expressivity.

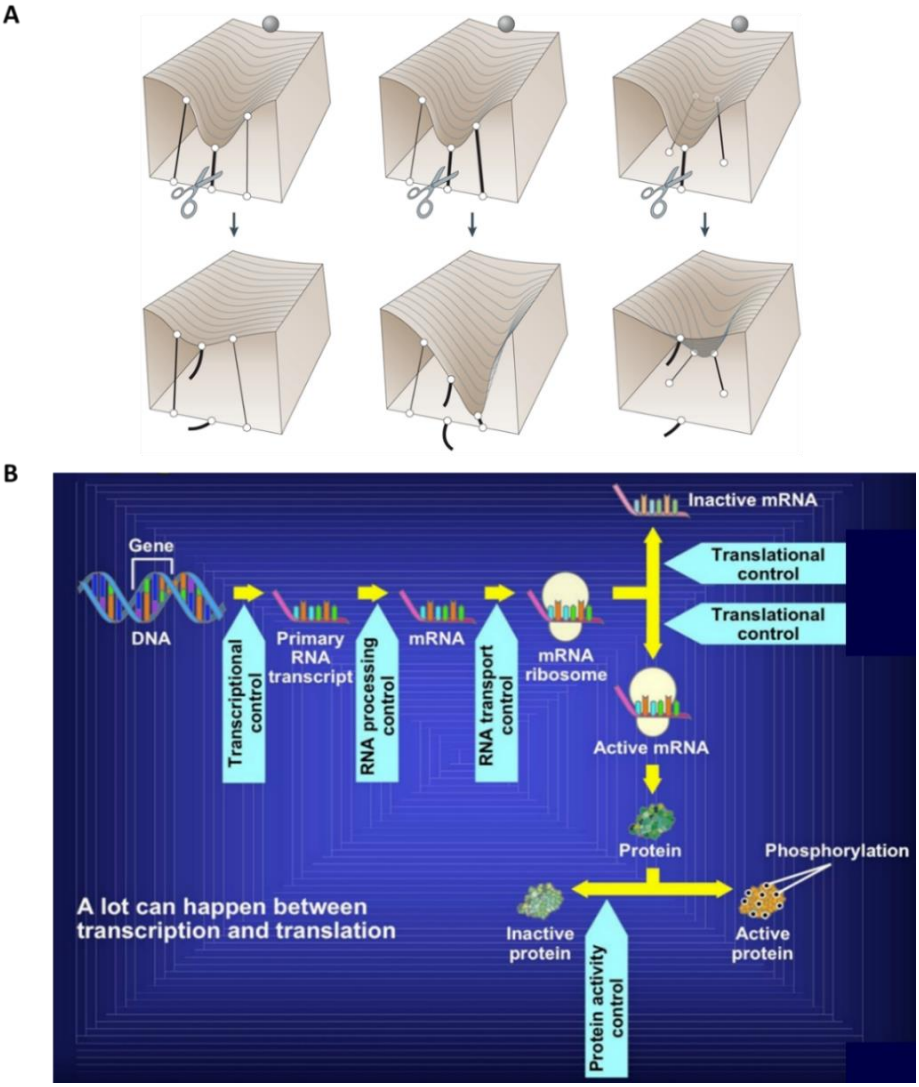


Figure 3.2 Genetic background effect and multiple steps affecting production of a protein from the gene. (A) The illustration represents how different genetic backgrounds can induce different phenotypic outcomes of a mutation (A modified "Waddington's epigenetic landscape" which is a metaphor for how gene regulation modulate development). A number of balls on the top of the hill roll down the hill. The balls will complete the journey through the grooves on the slope and end up to the lowest points. The different genetic backgrounds (represented as different hills controlled by underpinnings which make a certain shape of hills) vary at the molecular level (represented by guy-ropes of different thickness and configurations) but produce a consistent phenotype in wild type (The balls come to the same position after same journey). After disruption (depicted by breaking of the main rope, which represents a null mutation in a major gene), variation elsewhere reshapes the landscape which causes different destination of the balls depending on the hills (from Paaby and Rockman, 2014). (B) Image showing an overview of multiple steps influencing protein activity from gene expression to production of an active protein. A protein can affect the final protein product of another gene by interfering during the transcription and translation process. Six transitions along the gene expression pathway are marked as points where different controls may occur. Transcriptional control happens during the transcription of a gene into a primary RNA transcript. The second control, RNA processing control, occurs when a primary RNA transcript is processed into a mature mRNA transcript. The levels of mature RNA are also modulated by miRNAs. RNA transport control arises when the primary RNA transcript is transported to the ribosome. Translational controls occur during the transition of the ribosome-bound mRNA to either detaching from the ribosome, inactive state or an active state in which the mRNA remains bound to the ribosome. Then the active mRNA is translated into a protein molecule. The final control corresponding to protein activity regulation happens when the final protein product either remain inactivated or can be activated by for example phosphorylation (from Miko, 2008).

3.6 Effect of a mutation in function of the genetic background

In various model organisms such as bacteria (Remold and Lenski, 2004; Wang et al., 2013), yeast (Dowell et al., 2010), worm (Chandler, 2010; Paaby et al., 2015; Vu et al., 2015), Drosophila (Dworkin et al., 2009; Chari and Dworkin, 2013; Chandler et al., 2014), and mouse (Threadgill et al., 1995), studies have shown that the same mutation can lead to different phenotypes depending on the genetic background. For example, in Drosophila, a mutation of the *SCALLOPED (SD)* gene, encoding a transcription factor involved in wing determination, results in strong or mild effects on wing shape and size in function of the genetic background (Dworkin et al., 2009). The expression patterns of known *sd*-regulated genes showed organ-specific differential regulation depending on the genetic background. It was found that transcript expression profiles in wild types are significantly different compared to their mutants. Interestingly, the difference of transcript expression between the two wild types is as large as the effect of the mutation or even more pronounced. Furthermore, gene ontology categories were different between the genes enriched for the effect of a mutation of the *SCALLOPED (SD)* gene and the genes differentially expressed in two genetic backgrounds. It was suggested that both qualitative and quantitative differences at the level of downstream gene expression may cause genetic background-dependent mutation phenotype (Dworkin et al., 2009). To find interactors of sd^{E3} in different genetic backgrounds, a genome-wide modifier screen was conducted in the two backgrounds. This screen revealed that more than 70% of all modifiers of *sd* mutant phenotype behave in a genetic background dependent manner (Chari and Dworkin, 2013). Backcross mapping identified genomic regions contributing to the variation in the sd^{E3} phenotype using two genetic backgrounds showing opposite wing phenotype by sd^{E3}. Later, by conducting integrative genomic analysis with the data obtained previously, some candidate genes were identified that contribute to the genetic background effect of sd^{E3} (Chandler et al., 2014).

A recent study in worm where loss of function phenotypes were screened for 29 genes using RNA interference (RNAi) technology in 55 different genetic backgrounds revealed that wild type genomes contain numerous genetic modifiers which have little effect individually. However, the combination of those modifiers has dramatic influence in the penetrance of specific genes on complex traits such as embryogenesis (Paaby et al., 2015). Another study in worm showed that 20% of genes have different loss-of-function phenotypes in two different genetic backgrounds (Vu et al., 2015). Interestingly, the differences in severity of mutant phenotypes were predictable based on variation in the basal expression levels of the targeted gene in wild type. They showed that a lower expressed targeted gene in wild type leads to more severe phenotype in mutant and also influenced background dependent gene network by mutation could affect the phenotype (Figure 3.3A) (Vu et al., 2015).

In plants, genetic background effect on the penetrance of a mutation has been also analyzed. To study the evolution of self-incompatibility (SI), preventing self-pollination that

inhibits productive pollen-pistil interactions, in *Arabidopsis thaliana* which is self-fertile, transgenic lines were generated in different accessions. The two transgenes, *S-locus receptor kinase, SRK*, and *S-locus cysteine-rich protein, SCR*, that are required for self-incompatibility in *Arabidopsis lyrata* were introduced in 7 *Arabidopsis thaliana* accessions (Nasrallah et al., 2004). Only one accession (C24) showed a stable SI response among the seven accessions. This phenotype was not due to the level of stability or accumulation of the transcripts of the two introduced genes although variable expression levels were observed in the different accessions. Also polymorphisms in *S*-locus including *SRK* and *SCR* were found although these polymorphisms did not correlated with SI response. It was suggested that the SI response, such receptor-mediated signalling system, requires the formation of complexes between receptor, ligand, and possibly other proteins involving the activity of largely unknown elements of a downstream signalling cascade to have expected phenotype. Thus the polymorphism might affect to various components of the signalling complex or responsive pathway thereby it probably causes different phenotype.

A study from Gery et al., 2011 showed natural variation in freezing tolerance by using, in 8 accessions, RNAi-induced silencing of three *CBF* genes (*CBF1, 2*, and *3*) that play an important role in freezing tolerance. The degree of down-regulation of the *CBF* genes and the effect of this down-regulation were variable in the different accessions and in the different lines of the same accession. However, there is no simple correlation between *CBF* expression level and the freezing tolerance. This variable penetrance of the transgene could be explained by the fact that the independent evolution of natural populations might select different networks in different accessions related to stress response (Gery et al., 2011).

Although we have full genomic sequence of many model organisms, knowing how exactly gene-network for complex traits is working is still a challenge. The classic approach to study the dynamics of a network of interest is to perturb the network and then examine how it responds to the perturbation. The perturbation in a biological system can be obtained by changing the environment by external stimuli. For example, a plant can be exposed to cold/high temperature or different light conditions. A genetic perturbation in the network can also be obtained for one or more of the components of the network. Since natural variation provides a good source of genetic variation at multiple loci, combination of gene network study and natural variation allows us to understand how buffering of the network can produce different phenotypes (Figure 3.3B). The possibility to predict the penetrance of a mutation in different genetic backgrounds, which are genetically complex, has been questioned by researchers since it is critical for personalized medicine or targeted crop improvement. In this thesis, I tried to provide additional information on the natural variation response to genetic perturbation.

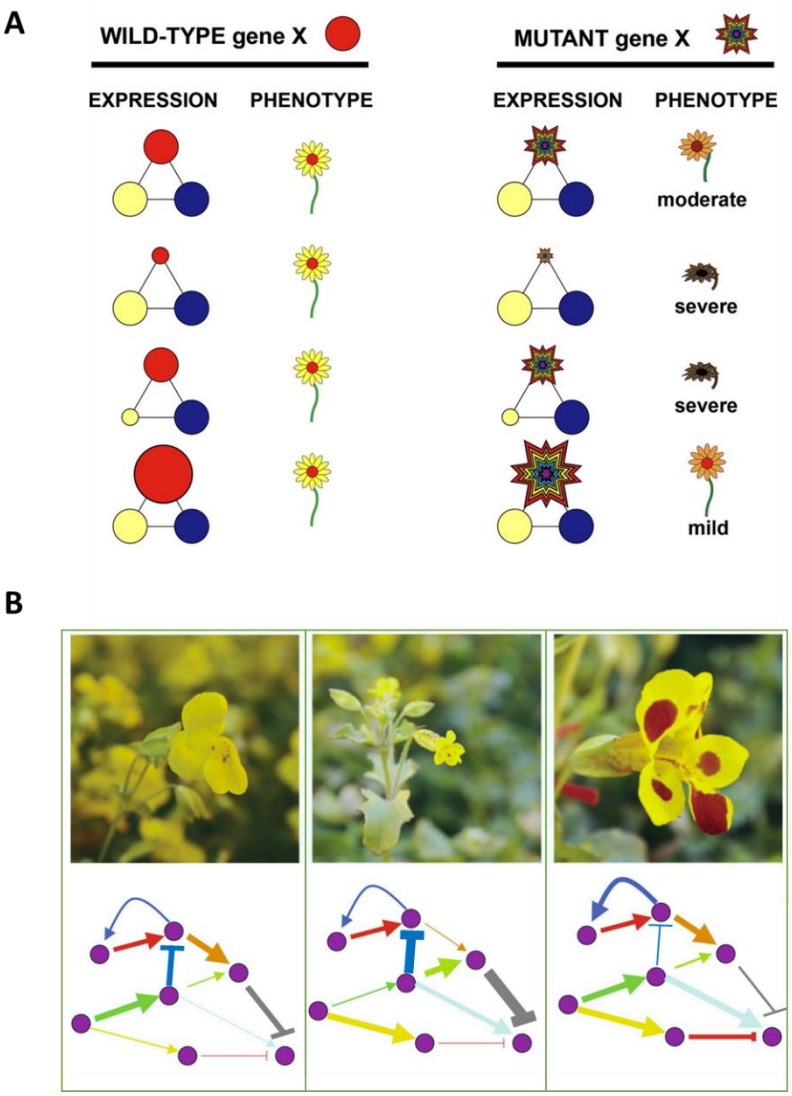


Figure 3.3 The effect of natural variation in gene expression in mutant phenotype and hypothetical network potentially determining flower shape and color between different genetic backgrounds. (A) Natural variation in gene expression leads to variation in the severity of mutant phenotypes. In the left part, natural variation in gene expression has little or no effect on phenotype in the context of wild-type alleles. In the right part, a mutation in a specific gene results in variation in expression pattern of itself. As a consequence, the gene networks are influenced in different manners depending on the genetic background. Thus, the phenotype of the mutation would be different as a result of genetic background specific gene network (from Vu et al., 2015). (B) The widespread species *M. guttatus* (left) has large, yellow flowers. In contrast, the flowers of *M. laciniatus* (center) are typically 75% smaller than those of *M. guttatus*. Other species have elevated expression of red anthocyanin pigments (right), as in this hybrid between subspecies of *M. luteus*. The network represented by differing widths of the connections [arrows] between network nodes [circles]) might be responsible for this natural variation (modified from Benfey and Mitchell-Olds, 2008).

References

Aarssen, L., and Clauss, M. (1992). Genotypic variation in fecundity allocation in Arabidopsis thaliana. Journal of Ecology**:** 109-114.

Alonso-Blanco, C., and Koornneef, M. (2000). Naturally occurring variation in Arabidopsis: an underexploited resource for plant genetics. Trends Plant Sci **5:** 22-29.

Alonso-Blanco, C., Aarts, M.G.M., Bentsink, L., Keurentjes, J.J.B., Reymond, M., Vreugdenhil, D., and Koornneef, M. (2009). What has natural variation taught us about plant development, physiology, and adaptation? Plant Cell **21:** 1877-1896.

Beemster, G.T., De Vusser, K., De Tavernier, E., De Bock, K., and Inze, D. (2002). Variation in growth rate between Arabidopsis ecotypes is correlated with cell division and A-type cyclin-dependent kinase activity. Plant Physiol **129:** 854-864.

Benfey, P.N., and Mitchell-Olds, T. (2008). From genotype to phenotype: systems biology meets natural variation. Science **320:** 495-497.

Boggs, N.A., Nasrallah, J.B., and Nasrallah, M.E. (2009). Independent S-locus mutations caused self-fertility in Arabidopsis thaliana. PLoS Genet **5:** e1000426.

Chandler, C.H. (2010). Cryptic intraspecific variation in sex determination in Caenorhabditis elegans revealed by mutations. Heredity (Edinb) **105:** 473-482.

Chandler, C.H., Chari, S., and Dworkin, I. (2013). Does your gene need a background check? How genetic background impacts the analysis of mutations, genes, and evolution. Trends Genet **29:** 358-366.

Chandler, C.H., Chari, S., Tack, D., and Dworkin, I. (2014). Causes and consequences of genetic background effects illuminated by integrative genomic analysis. Genetics **196:** 1321-1336.

Chari, S., and Dworkin, I. (2013). The conditional nature of genetic interactions: the consequences of wild-type backgrounds on mutational interactions in a genome-wide modifier screen. PLoS Genet **9:** e1003661.

Clauw, P., Coppens, F., De Beuf, K., Dhondt, S., Van Daele, T., Maleux, K., Storme, V., Clement, L., Gonzalez, N., and Inze, D. (2015). Leaf responses to mild drought stress in natural variants of Arabidopsis. Plant Physiol **167:** 800-816.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J **16:** 735-743.

Cronk, Q.C. (2001). Plant evolution and development in a post-genomic context. Nat Rev Genet **2:** 607-619.

de Montaigu, A., Giakountis, A., Rubin, M., Toth, R., Cremer, F., Sokolova, V., Porri, A., Reymond, M., Weinig, C., and Coupland, G. (2015). Natural diversity in daily rhythms of gene expression contributes to phenotypic variation. Proc Natl Acad Sci U S A **112:** 905-910.

Delker, C., Pöschl, Y., Raschke, A., Ullrich, K., Ettingshausen, S., Hauptmann, V., Grosse, I., and Quint, M. (2010). Natural variation of transcriptional auxin response networks in *Arabidopsis thaliana*. Plant Cell **22:** 2184-2200.

Dhondt, S., Wuyts, N., and Inzé, D. (2013). Cell to whole-plant phenotyping: the best is yet to come. Trends Plant Sci. **18:** 428-439.

Dhondt, S., Gonzalez, N., Blomme, J., De Milde, L., Van Daele, T., Van Akoleyen, D., Storme, V., Coppens, F., Beemster, G.T.S., and Inzé, D. (2014). High-resolution time-resolved imaging of *in vitro* Arabidopsis rosette growth. Plant J. **80:** 172-184.

Diamond, J. (2002). Evolution, consequences and future of plant and animal domestication. Nature **418:** 700-707.

Dowell, R.D., et al. (2010). Genotype to phenotype: a complex problem. Science **328:** 469.

Dworkin, I., Kennerly, E., Tack, D., Hutchinson, J., Brown, J., Mahaffey, J., and Gibson, G. (2009). Genomic consequences of background effects on scalloped mutant expressivity in the wing of Drosophila melanogaster. Genetics **181:** 1065-1076.

El-Lithy, M.E., Clerkx, E.J., Ruys, G.J., Koornneef, M., and Vreugdenhil, D. (2004). Quantitative trait locus analysis of growth-related traits in a new Arabidopsis recombinant inbred population. Plant Physiol **135:** 444-458.

Gery, C., Zuther, E., Schulz, E., Legoupi, J., Chauveau, A., McKhann, H., Hincha, D.K., and Teoule, E. (2011). Natural variation in the freezing tolerance of Arabidopsis thaliana: effects of RNAi-induced CBF depletion and QTL localisation vary among accessions. Plant Sci **180:** 12-23.

Huang, X., and Han, B. (2014). Natural variations and genome-wide association studies in crop plants. Annu. Rev. Plant Biol. **65:** 531-551.

Huang, X., Effgen, S., Meyer, R.C., Theres, K., and Koornneef, M. (2012). Epistatic natural allelic variation reveals a function of AGAMOUS-LIKE6 in axillary bud formation in Arabidopsis. Plant Cell **24:** 2364-2379.

Huang, X., Paulo, M.J., Boer, M., Effgen, S., Keizer, P., Koornneef, M., and van Eeuwijk, F.A. (2011). Analysis of natural allelic variation in Arabidopsis using a multiparent recombinant inbred line population. Proc Natl Acad Sci U S A **108:** 4488-4493.

Keurentjes, J.J., Fu, J., Terpstra, I.R., Garcia, J.M., van den Ackerveken, G., Snoek, L.B., Peeters, A.J., Vreugdenhil, D., Koornneef, M., and Jansen, R.C. (2007). Regulatory network construction in Arabidopsis by using genome-wide gene expression quantitative trait loci. Proc Natl Acad Sci U S A **104:** 1708-1713.

Kover, P.X., Valdar, W., Trakalo, J., Scarcelli, N., Ehrenreich, I.M., Purugganan, M.D., Durrant, C., and Mott, R. (2009). A Multiparent Advanced Generation Inter-Cross to fine-map quantitative traits in Arabidopsis thaliana. PLoS Genet **5:** e1000551.

Kramer, U. (2015). Planting molecular functions in an ecological context with Arabidopsis thaliana. eLife **4**.

Li, B., Suzuki, J.-I., and Hara, T. (1998). Latitudinal variation in plant size and relative growth rate in Arabidopsis thaliana. Oecologia **115:** 293-301.

Loudet, O., Chaillou, S., Merigout, P., Talbotec, J., and Daniel-Vedele, F. (2003). Quantitative trait loci analysis of nitrogen use efficiency in Arabidopsis. Plant Physiol. **131:** 345-358.

Maloof, J.N., Borevitz, J.O., Dabi, T., Lutes, J., Nehring, R.B., Redfern, J.L., Trainer, G.T., Wilson, J.M., Asami, T., Berry, C.C., Weigel, D., and Chory, J. (2001). Natural variation in light sensitivity of Arabidopsis. Nat Genet **29:** 441-446.

Massonnet, C., Tisné, S., Radziejwoski, A., Vile, D., De Veylder, L., Dauzat, M., and Granier, C. (2011). New insights into the control of endoreduplication: endoreduplication could be driven by organ growth in Arabidopsis leaves. Plant Physiol. **157:** 2044-2055.

Meijón, M., Satbhai, S.B., Tsuchimatsu, T., and Busch, W. (2014). Genome-wide association study using cellular traits identifies a new regulator of root development in *Arabidopsis*. Nat. Genet. **46:** 77-81.

Meyer, R.C., Steinfath, M., Lisec, J., Becher, M., Witucka-Wall, H., Torjek, O., Fiehn, O., Eckardt, A., Willmitzer, L., Selbig, J., and Altmann, T. (2007). The metabolic signature related to high plant growth rate in Arabidopsis thaliana. Proc Natl Acad Sci U S A **104:** 4759-4764.

Miko, I. (2008). Phenotype variability: penetrance and expressivity. Nature Education **1:** 137.

Miles, C., and Wayne, M. (2008). Quantitative trait locus (QTL) analysis. Nature Education **1:** 208.

Millenaar, F.F., Cox, M.C., van Berkel, Y.E., Welschen, R.A., Pierik, R., Voesenek, L.A., and Peeters, A.J. (2005). Ethylene-induced differential growth of petioles in Arabidopsis. Analyzing natural variation, response kinetics, and regulation. Plant Physiol **137:** 998-1008.

Miller, M., Song, Q., Shi, X., Juenger, T.E., and Chen, Z.J. (2015). Natural variation in timing of stress-responsive gene expression predicts heterosis in intraspecific hybrids of Arabidopsis. Nat Commun **6:** 7453.

Nasrallah, M.E., Liu, P., Sherman-Broyles, S., Boggs, N.A., and Nasrallah, J.B. (2004). Natural variation in expression of self-incompatibility in Arabidopsis thaliana: implications for the evolution of selfing. Proc Natl Acad Sci U S A **101:** 16070-16074.

Paaby, A.B., and Rockman, M.V. (2014). Cryptic genetic variation: evolution's hidden substrate. Nat Rev Genet **15:** 247-258.

Paaby, A.B., White, A.G., Riccardi, D.D., Gunsalus, K.C., Piano, F., and Rockman, M.V. (2015). Wild worm embryogenesis harbors ubiquitous polygenic modifier variation. eLife **4**.

Pérez-Pérez, J.M., Serrano-Cartagena, J., and Micol, J.L. (2002). Genetic analysis of natural variations in the architecture of *Arabidopsis thaliana* vegetative leaves. Genetics **162:** 893-915.

Pigliucci, M., and Schlichting, C.D. (1995). Reaction norms of Arabidopsis (Brassicaceae). III. Response to nutrients in 26 populations from a worldwide collection. Am. J. Bot.**:** 1117-1125.

Reiter, R.S., Williams, J.G., Feldmann, K.A., Rafalski, J.A., Tingey, S.V., and Scolnik, P.A. (1992). Global and local genome mapping in Arabidopsis thaliana by using recombinant inbred lines and random amplified polymorphic DNAs. Proc Natl Acad Sci U S A **89:** 1477-1481.

Remold, S.K., and Lenski, R.E. (2004). Pervasive joint influence of epistasis and plasticity on mutational effects in Escherichia coli. Nat Genet **36:** 423-426.

Schadt, E.E., et al. (2003). Genetics of gene expression surveyed in maize, mouse and man. Nature **422:** 297-302.

Sulpice, R., Nikoloski, Z., Tschoep, H., Antonio, C., Kleessen, S., Larhlimi, A., Selbig, J., Ishihara, H., Gibon, Y., Fernie, A.R., and Stitt, M. (2013). Impact of the carbon and nitrogen supply on relationships and connectivity between metabolism and biomass in a broad panel of Arabidopsis accessions. Plant Physiol **162:** 347-363.

Sulpice, R., et al. (2010). Network analysis of enzyme activities and metabolite levels and their relationship to biomass in a large panel of Arabidopsis accessions. Plant Cell **22:** 2872-2893.

Sulpice, R., et al. (2009). Starch as a major integrator in the regulation of plant growth. Proc Natl Acad Sci U S A **106:** 10348-10353.

Sureshkumar, S., Todesco, M., Schneeberger, K., Harilal, R., Balasubramanian, S., and Weigel, D. (2009). A genetic defect caused by a triplet repeat expansion in Arabidopsis thaliana. Science **323:** 1060-1063.

Threadgill, D.W., Dlugosz, A.A., Hansen, L.A., Tennenbaum, T., Lichti, U., Yee, D., LaMantia, C., Mourton, T., Herrup, K., Harris, R.C., and et al. (1995). Targeted disruption of mouse EGF receptor: effect of genetic background on mutant phenotype. Science **269:** 230-234.

Turesson, G. (1922). The genotypical response of the plant species to the habitat. Hereditas **3:** 211-350.

van Leeuwen, H., Kliebenstein, D.J., West, M.A.L., Kim, K., van Poecke, R., Katagiri, F., Michelmore, R.W., Doerge, R.W., and St. Clair, D.A. (2007). Natural variation among *Arabidopsis thaliana* accessions for transcriptome response to exogenous salicylic acid. Plant Cell **19:** 2099-2110.

Vu, V., Verster, A.J., Schertzberg, M., Chuluunbaatar, T., Spensley, M., Pajkic, D., Hart, G.T., Moffat, J., and Fraser, A.G. (2015). Natural Variation in Gene Expression Modulates the Severity of Mutant Phenotypes. Cell **162:** 391-402.

Wang, Q., Sajja, U., Rosloski, S., Humphrey, T., Kim, M.C., Bomblies, K., Weigel, D., and Grbic, V. (2007). HUA2 caused natural variation in shoot morphology of A. thaliana. Current biology : CB **17:** 1513-1519.

Wang, Y., Arenas, C.D., Stoebel, D.M., and Cooper, T.F. (2013). Genetic background affects epistatic interactions between two beneficial mutations. Biol Lett **9:** 20120328.

Weigel, D. (2012). Natural variation in Arabidopsis: from molecular genetics to ecological genomics. Plant Physiol **158:** 2-22.

West, M.A., Kim, K., Kliebenstein, D.J., van Leeuwen, H., Michelmore, R.W., Doerge, R.W., and St Clair, D.A. (2007). Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in Arabidopsis. Genetics **175:** 1441-1450.

Chapter 4. Scope and outline of the thesis

Leaf growth and its quantitative output, final size, are the result of two different processes, cell division and cell expansion which are regulated both by environmental and genetic factors. A number of growth-promoting genes and networks have been identified that regulate leaf growth and biomass in Arabidopsis (Gonzalez et al., 2012; Hepworth and Lenhard, 2014). To functionally analyze genes involved in the regulation of leaf growth, different approaches can be used. Particularly, artificially generated genetic perturbation has been commonly used to investigate the function of genes of interest (Blomme et al., 2014; Vanhaeren et al., 2014). Gene inactivation allows for studying gene function in its naturally occurring genetic environment whereas ectopic overexpression of genes allows for analyzing the consequences of disturbing the normal spatial and temporal expression of the genes of interest. Such genetic perturbations are most often carried out in one genetic background. For example in Arabidopsis, genetic perturbations are mainly performed in the Columbia-0 background (Col-0). However, mutations and transgenes possibly behave very differently in different genetic backgrounds (Chandler et al., 2013). Although the influence of genetic background on the output of a mutation or transgene was observed in several model organisms (Chandler et al., 2013), the knowledge about this phenomenon in plants is poorly investigated.

Analysis of the response of genetic perturbations in 17 Arabidopsis accessions

The accurate prediction of background-dependent phenotypic effects of a specific genetic perturbation is of great interest for crop breeders in order to improve crops with a stable effect of genetic perturbations across different genetic backgrounds.

To understand how the genetic interactions between transgenes and genetic backgrounds affect the phenotypic output in Arabidopsis, we chose three genes known to promote leaf growth in Col-0 background when their expression is modified: overexpression of *GA20ox1* (*GA20ox1*^{OE}) (Gonzalez et al., 2010), overexpression of artificial microRNA *targeting PPD (amiPPD)* (Gonzalez et al., 2015), and overexpression of dominant negative form of *DA1 (DN-DA1*^{OE}*)* (Li et al., 2008) (Figure 4.1). We chose these genes because they are regulating different processes during leaf development. Both *DA1* and *GA20ox1* are involved in the regulation of cell proliferation and *PPD* has a function during meristemoid division. *GA20ox1* is also involved in the regulation of the cell expansion phase. Besides, these three genes were selected since they encode proteins that have different biological functions: GA20ox1 is a rate-limiting gibberellin biosynthesis enzyme, PPD is a transcription regulator, and DA1 is an ubiquitin receptor (see Chapter 2 for details). The three different transgenes were introduced in seventeen Arabidopsis natural accessions from around the world including Col-0 (Figure 3.1 and Table 4.1). Originally, seventeen Arabidopsis accessions were selected to cover most common

genetic variability of Arabidopsis thaliana. Since some accessions were problematic to obtain transformants, six additional accessions were included based on previous experiments done in the System Biology of yield group (Clauw et al., 2015). Two commonly used accessions (C24 and Ler-0) were also included in this project.

In this project, we used two major methods: phenotypic measurements and RNA sequencing analysis from the transgenics generated. By using these approaches, we set out to answer three main research questions.

(1) From the phenotypic data, what is the variation of the penetrance of the transgene expression in different genetic backgrounds?

(2) From the RNA sequencing data:

- Which processes are differentially regulated compared to wild type and are they similar?
- How can the difference in penetrance be molecularly explained?

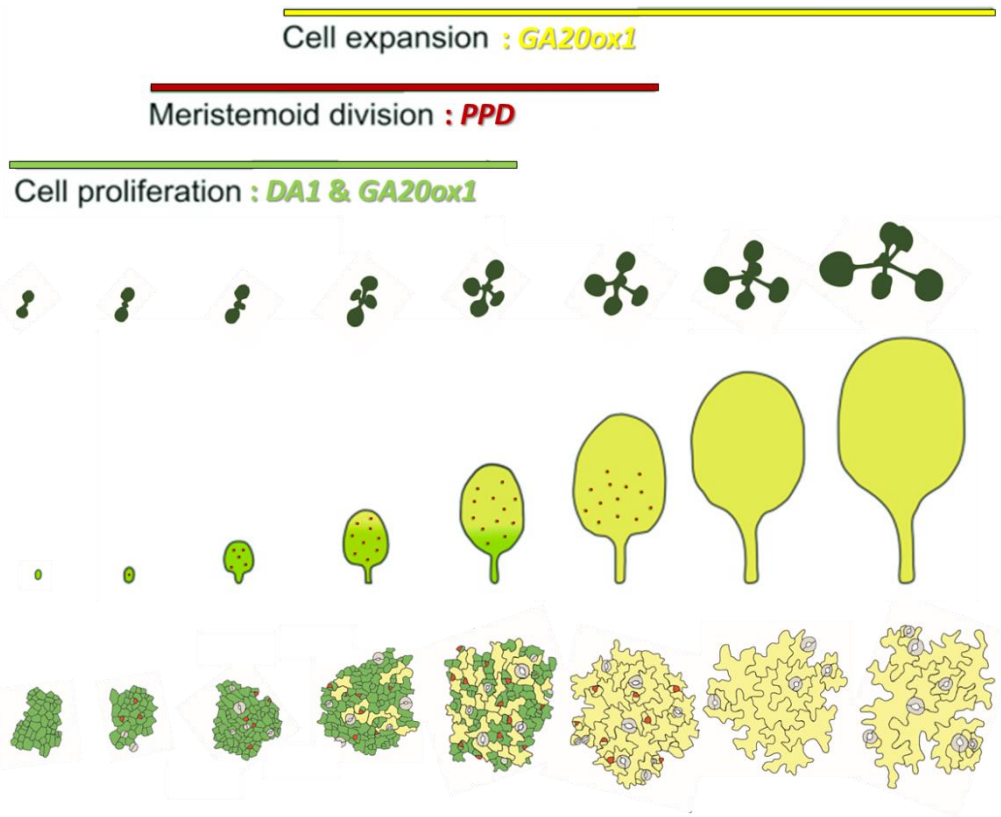


Figure 4.1 The three genes (*GA20ox1*, *PPD*, and *DA1*) studied in this project are involved in different developmental processes during leaf development in Arabidopsis. The three genes play roles in different phases of leaf development (cell proliferation; *DA1* and *GA20ox1*, meristemoid division; *PPD* and cell expansion; *GA20ox1*). Images of rosette, leaf 1 and 2, and cell drawing overtime are shown. Dividing cells in the primary general cell-division front are represented in green, meristemoid cells in red and expanding cells in yellow. Cell drawings are from the abaxial side of leaf epidermis (modified from Gonzalez et al., 2012; Vanhaeren et al., 2015).

Table 4.1 Geographic origin of the 17 Arabidopsis accessions used in this study. CS; Arabidopsis biological resource center.

Accession	Origin	CS stock number
An-1	Belgium	CS76091
Blh-1	Czech Republic	CS76098
C24	Portugal	CS76106
Col-0	Poland	CS76113
Cvi-0	Cape Verdi	CS76116
Ey15-2	Germany	CS76399
ICE138	Central Asia	CS76426
ICE153	Central Asia	CS76381
ICE163	Southern Tyrol	CS76353
ICE61	Russia	CS76378
ICE75	Russia	CS76422
ICE97	Southern Italy	CS76359
Ler-0	Germany	CS77020
Oy-0	Norway	CS76203
Sha	Tadjikistan	CS76382
WalhaesB4	Germany	CS76408
Yeg-1	Kaukasus	CS76394

Experimental design for the analysis

The sequence of the genes from Col-0 were cloned and introduced into 17 accessions. At least one and maximum five T_3 homozygote lines having a single locus insertion site of the transgene in each accession have been selected for further analysis (Table 4.2). The selection of the transgenic lines for further analysis was done based on the expression levels of the (trans)gene in T_2 generation. For the overexpression lines, *GA20ox1*^{OE} and *DN-DA1*^{OE}, maximum five of "the highest overexpressing" lines still showing a range of expression were chosen and for the *amiPPD* lines, maximum five of "the lowest *PPD* expressing" lines as well still showing a range of expression were selected. Since some transgenic lines were problematic to set flowers or make enough seeds and to get homozygote lines, not always five independent lines were generated for further analysis (Table 4.2).

Using an in house phenotyping system called MIRGIS (multi-camera in vivo rosette growth imaging system), rosette growth and leaf size were analysed for all transgenic lines with their corresponding wild type (Figure 4.2B). Plants were grown for 25 days after stratification (DAS) and a picture was taken every day to monitor and analyse rosette growth. At 25 DAS, leaves were harvested for measuring leaf size-related parameters at rosette, leaf, and cellular levels (Figure 4.2C). For *GA20ox1*-overexpression and *amiPPD* lines, proliferating leaf 6 and leaf 3 were harvested to conduct RNA sequencing analysis (see Chapter 5 and Chapter 6), respectively (Figure 4.2A). Further correlation analysis was done based on these phenotypic and molecular data.

Here, I present an overview of the structure of my thesis.

In the **first part**, I presented background knowledge of leaf growth in Arabidopsis (**Chapter 1**) and the three genes studied in this project (**Chapter 2**). In the last chapter of introduction, I described the use of natural variation in plant research (**Chapter 3**).

The **second part** of this thesis contains four research chapters that detail the major findings of this project. In **Chapter 5**, we analyze natural variation in response to gibberellin changes in Arabidopsis by using overexpression of *GA20ox1* in 17 different Arabidopsis accessions. Phenotypic variability of the 17 wild-type accessions is first analyzed and the penetrance and expressivity of *GA20ox1*^{OE} as well as gibberellin contents and transcriptome response in the different accessions are the studied.

In **Chapter 6**, morphological and molecular variation in response to the down-regulation of *PEAPOD* is examined in 15 Arabidopsis accessions.

In **Chapter 7**, I provide data on the phenotypic variation in 17 Arabidopsis accessions caused by overexpression of a dominant negative form of *DA1*. In this chapter, I also combined all phenotypic data of the transgenic lines from Chapter 5, 6, and 7 and compared the effects of the three transgenes in 17 accessions.

In the last research chapter, **Chapter 8**, I perform comparative analysis between destructive (leaf series data) and non-destructive methods (MIRGIS data) for measuring rosette area by using the data from the transgenic lines in 17 accessions we generated in this study.

In the last, **third part** of this thesis, I give an overview of the main results of the different chapters and present a broader general discussion. Additionally, future perspectives and methods to address the inevitable new questions generated by this study are discussed.

Table 4.2 The number of homozygote transgenic lines obtained for further analysis per accession and per transgenic lines.

	***GA20ox1*^{OE}**	***amiPPD***	***DN-DA1*^{OE}**
An-1	5	5	5
Blh-1	5	3	4
C24	5	5	3
Col-0	3	5	5
Cvi-0	3	2	2
Ey15-2	5	5	4
ICE138	2	3	2
ICE153	4	4	2
ICE163	5	2	1
ICE61	3	5	4
ICE75	4	/	2
ICE97	3	/	1
Ler-0	5	4	4
Oy-0	4	5	4
Sha	3	3	5
WalhaesB4	4	2	3
Yeg-1	2	3	1

/; no transgenics

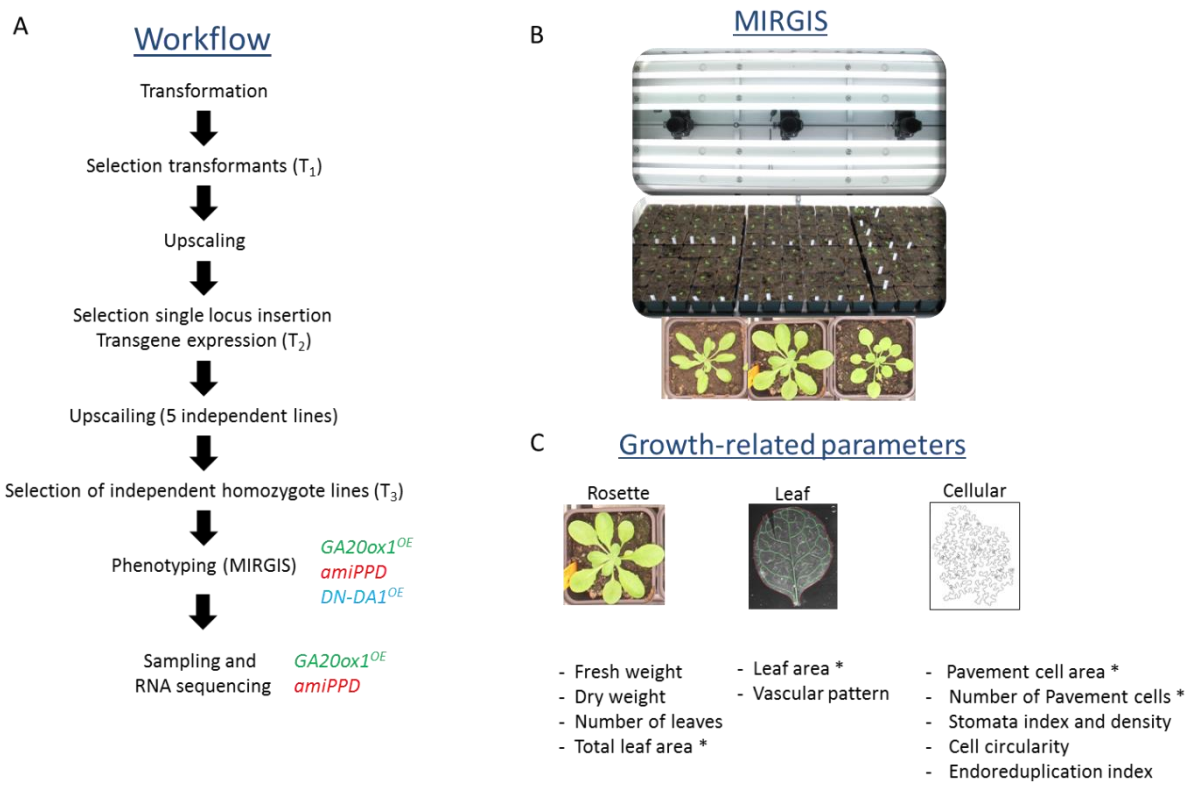


Figure 4.2 Overview of the work scheme of this project. (A) Workflow of this project from the transformation of the plants to the phenotyping with 3 different transgenes (*GA20ox1-overexpression* (OE), *amiPPD*, and overexpression of dominant negative form of *DA1*). After selecting homozygote lines, all transgenic lines were phenotyped in the MIRGIS (B) platform. (C) Several growth-related parameters at different levels (rosette, leaf, and cellular levels) were measured in only the 17 wild type accessions and some of these parameters (marked with *) in the transgenic lines. For RNA sequencing analysis were done in transgenic lines of *GA20ox1*^{OE} and *amiPPD*.

References

Blomme, J., Inze, D., and Gonzalez, N. (2014). The cell-cycle interactome: a source of growth regulators? J Exp Bot **65:** 2715-2730.

Chandler, C.H., Chari, S., and Dworkin, I. (2013). Does your gene need a background check? How genetic background impacts the analysis of mutations, genes, and evolution. Trends Genet **29:** 358-366.

Clauw, P., Coppens, F., De Beuf, K., Dhondt, S., Van Daele, T., Maleux, K., Storme, V., Clement, L., Gonzalez, N., and Inze, D. (2015). Leaf responses to mild drought stress in natural variants of Arabidopsis. Plant Physiol **167:** 800-816.

Gonzalez, N., Vanhaeren, H., and Inzé, D. (2012). Leaf size control: complex coordination of cell division and expansion. Trends Plant Sci. **17:** 332-340.

Gonzalez, N., et al. (2010). Increased leaf size: different means to an end. Plant Physiol. **153:** 1261-1279.

Gonzalez, N., et al. (2015). A Repressor Protein Complex Regulates Leaf Growth in Arabidopsis. Plant Cell **27:** 2273-2287.

Hepworth, J., and Lenhard, M. (2014). Regulation of plant lateral-organ growth by modulating cell number and size. Curr Opin Plant Biol **17:** 36-42.

Li, Y., Zheng, L., Corke, F., Smith, C., and Bevan, M.W. (2008). Control of final seed and organ size by the *DA1* gene family in *Arabidopsis thaliana*. Genes Dev. **22:** 1331-1336.

Vanhaeren, H., Gonzalez, N., and Inze, D. (2015). A Journey Through a Leaf: Phenomics Analysis of Leaf Growth in Arabidopsis thaliana. Arabidopsis Book **13:** e0181.

Vanhaeren, H., Gonzalez, N., Coppens, F., De Milde, L., Van Daele, T., Vermeersch, M., Eloy, N.B., Storme, V., and Inze, D. (2014). Combining growth-promoting genes leads to positive epistasis in Arabidopsis thaliana. eLife **3:** e02252.

Part 2. Results

Chapter 5

Natural variation of gibberellin responses in *Arabidopsis thaliana*

Youn-Jeong Nam[1,2], Dorota Herman[1,2], Eunyoung Chae[3], Kojima, Mikiko[4], Frederik Coppens[1,2], Veronique Storme[1,2], Twiggy Van Daele[1,2], Sakakibara, Hitoshi[4], Detlef Weigel[3], Dirk Inzé[1,2*] & Nathalie Gonzalez[1,2*]

[1]Department of Plant Systems Biology, VIB B-9052 Gent, Belgium, [2]Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Gent, Belgium, [3]Department of Molecular Biology, Max Planck Institute for Developmental Biology, Tübingen, Germany, [4]RIKEN Center for Sustainable Resource Science, 1-7-22 Suehiro, Tsurumi, Yokohama, Kanagawa 230-0045, Japan.
* These authors contributed equally to this work.

Contributions: Y.J.N. was the main author of this work. Y.J.N., E.C, K.M, and T.V.D. conducted experimental work. D.H., F.C., V.S., and Y.J.N. analysed the data. S.H., D.I., N.G., and D.W. supervised the research. D.W., D.I. and N.G. contributed to the writing of this chapter.

This chapter has been summited to Genome biology.

Introduction

The relationship between a phenotype and a specific genetic change, also referred to as expressivity, depends not only on the environment, but also on the genetic background in which a mutation occurs (Dowell et al., 2010; Chandler et al., 2013; Chari and Dworkin, 2013). Although typically treated as a nuisance by laboratory geneticists, such epistatic interactions are not only central to studies of genetic variation in populations but can also inform our understanding of genetic networks and phenotypic robustness (Félix, 2007; Félix and Wagner, 2008; Paaby et al., 2015; Vu et al., 2015). Similar to its implications for human health (Schilsky, 2010), the accurate prediction of background-dependent phenotypic effects of specific mutations is of great interest to crop breeders.

Gibberellins (GAs) are phytohormones with well-documented roles in germination, stem elongation, flowering, and leaf-, seed- and fruit development, often in response to environmental changes (Hedden, 2003; Ueguchi-Tanaka et al., 2007; Schwechheimer and Willige, 2009; Claeys et al., 2014). In addition, roles in plant immunity have been discovered (De Bruyne et al., 2014). GA20-oxidase (GA20ox), a rate-limiting enzyme in the GA biosynthesis pathway, catalyses consecutive oxidation events in the late steps of the formation of active GAs. It uses various intermediates as substrates, including GA_{12}, GA_{53}, GA_{15}, GA_{44}, GA_{24}, and GA_{19}, to form as a final step GA_9 and/or GA_{20}, two substrates of GA3-oxidase (GA3ox), that are then converted into bioactive GAs (Hedden and Thomas, 2012). Overexpression of *GA20ox* has been shown to enhance plant growth as a result of increased GA levels (Huang et al., 1998; Coles et al., 1999; Gonzalez et al., 2010; Nelissen et al., 2012).

Here, to assess natural variation in the ability to respond to changes in gibberellin metabolism, we examined at multiple levels the effect of the ectopic expression of *GA20ox1* in 17 *Arabidopsis thaliana* accessions. Our results indicate that hormone metabolism and signalling are remarkably different in these accessions, supporting a potential role of hormone responses in adaptation to the environment.

Results

Natural variation in growth and hormone content

Seventeen accessions from throughout the native range of the species (Table 4.1) were grown for 25 days after stratification (DAS) in soil. Thirteen leaf size-related parameters were measured at rosette (fresh and dry weight, number of leaves, total rosette area), leaf (first leaf pair area, vascular complexity and density), and cellular level (stomatal index and density, epidermal pavement cell number, area, and circularity, and endoreduplication index of the first leaf pair).

The 17 accessions, which included the reference accession Col-0, varied for all parameters (Figure 5.1A). They differed more than 2.5 fold in rosette biomass, total rosette area,

pavement cell number and area, stomatal density and vascular complexity. Fresh weight significantly positively correlated with total rosette area, leaf number, and leaf area and negatively with vascular density and complexity (Extended Data Figure 5.1).

To examine how much of growth variation could be explained by differences in the phytohormone accumulation, we measured biosynthetic intermediates and different bio-active forms of GA, cytokinin, salicylic acid (SA), jasmonic acid (JA), abscisic acid (ABA), and auxin in 12 DAS seedlings (Extended Data Table 5.1).

GAs, SA and the auxin IAA varied the most, while cytokinins and ABA varied the least, with JA showing an intermediate degree of changes (Figure 5.1B and Figure 5.1C). In addition, we found that the relationships between different GAs and their intermediates, most of which are substrates of GA20ox, were complex. For example, the bioactive GA_4, showed a similar profile as its direct precursor, GA_9. However, concentrations of all the other intermediates that we measured did not parallel GA_9 and GA_4 levels, suggesting utilization of GA20ox activity in the GA biosynthesis differ in accessions (Figure 5.1B). Several hormones had mostly significant positive correlation with each other (Extended Data Table 5.1 and Extended Data Figure 5.2).

We uncovered that three hormones, GA, iP and IAA, are significantly positively correlated with pavement cell number, a leaf-growth parameter (Extended Data Figure 5.3). Furthermore, one of the GA20ox products, GA_{19} and the GA bio-inactive form, GA_8, were negatively correlated with the other two leaf-growth parameters, endoreduplication index and stomatal index.

Consequences of *GA20ox1* overexpression in different accessions

Overexpression of *GA20ox1* in the reference Col-0 background causes similar phenotypes as treatment with exogenous GA, such as larger rosette leaves, longer hypocotyls, increased height and early flowering (Huang et al., 1998; Coles et al., 1999; Gonzalez et al., 2010; Nelissen et al., 2012; Ribeiro et al., 2012). To investigate natural genetic variation in phenotypic responses to GA level perturbance in *A. thaliana*, we introduced the same overexpression construct into 16 additional accessions. Two to five independent homozygous lines for each accession were selected based on the expression levels of the *GA20ox1* in T_2 generation. Maximum five of "the highest overexpressing" lines, still showing a range of expression, were chosen and homozygous lines were obtained for further analysis. The homozygous lines were grown in soil for 25 days and leaf area was measured (Extended Data Table 5.2). Most of the independent transgenic lines within an accession showed consistent response to the transgene. However, there were some exceptions for instance, line 2 of WalhaesB4 and line 1 of ICE97 showed somewhat different effect compared to the other independent lines.

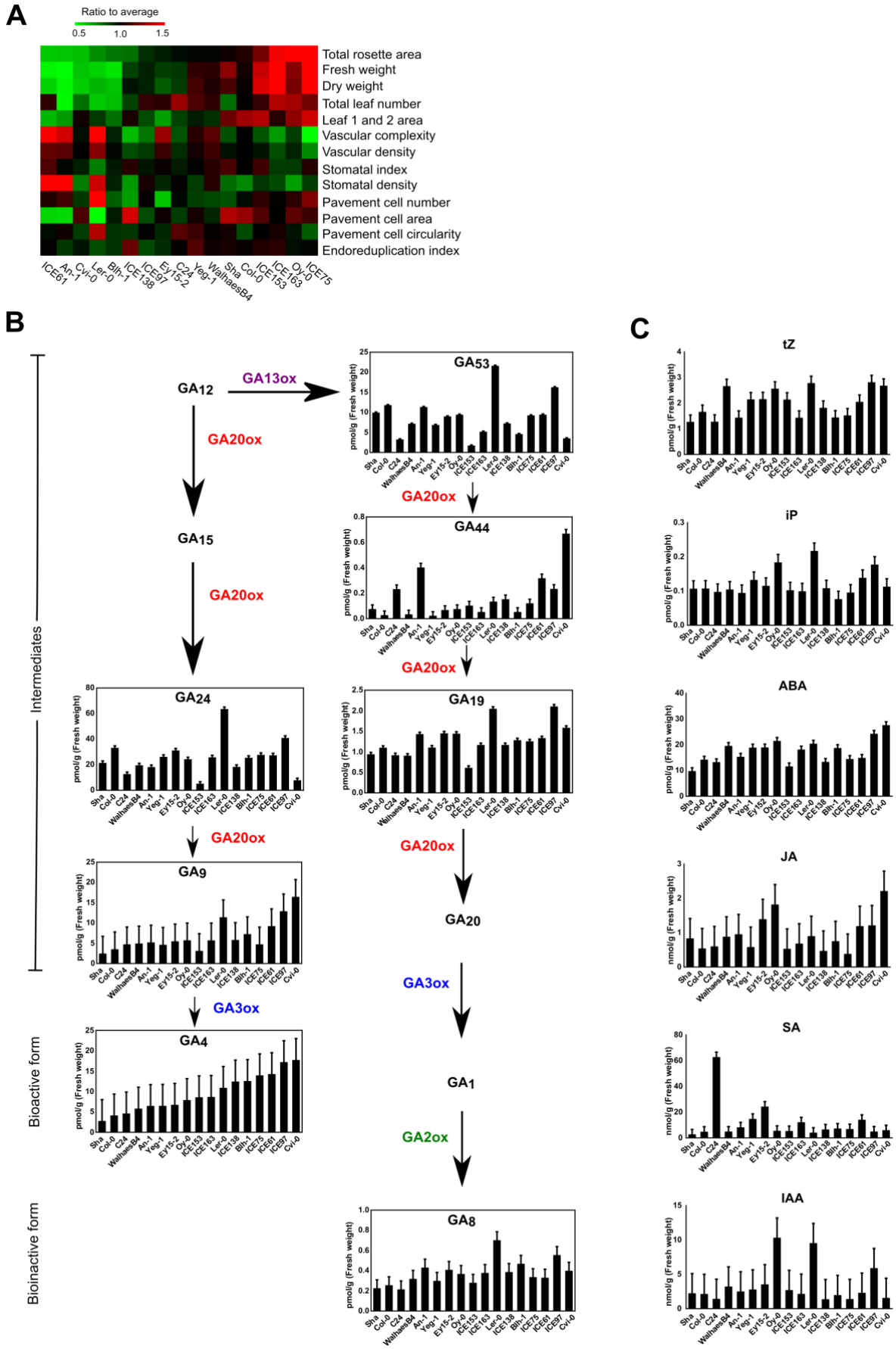

A
Ratio to average
0.5
1.0
1.5
Total rosette area
Fresh weight
Dry weight
Total leaf number
Leaf 1 and 2 area
Vascular complexity
Vascular density
Stomatal index
Stomatal density
Pavement cell number
Pavement cell area
Pavement cell circularity
Endoreduplication index
ICE61
An-1
Cvi-0
Ler-0
Blh-1
ICE138
ICE97
Ey15-2
C24
Yeg-1
WalhaesB4
Sha
Col-0
ICE153
ICE163
Oy-0
ICE75
B
Intermediates
Bioactive form
Bioinactive form
GA12
GA13ox
GA53
GA20ox
GA15
GA20ox
GA44
GA20ox
GA24
GA20ox
GA19
GA20ox
GA9
GA20
GA3ox
GA4
GA3ox
GA1
GA2ox
GA8
pmol/g (Fresh weight)
C
tZ
iP
ABA
JA
SA
IAA
pmol/g (Fresh weight)
nmol/g (Fresh weight)

Figure 5.1 Variability in leaf size-related parameters and hormone content in 17 Arabidopsis accessions. (A) Heat map representing the distance to the average of 17 accessions for 13 leaf size-related parameters (N=3). Accessions are arranged based on the value of the rosette area. (B) Basal GA levels in 17 accessions. GA biosynthesis (GA20ox and GA3ox) and catabolic (GA2ox) enzymes are indicated with different colors. Levels of GA_{20} and GA_1 were not detected. (C) Basal levels of cytokinins (tZ and iP), ABA, JA, SA, and IAA in the 17 accessions (N=3). Error bars represent standard error.

Most, but not all, accessions visibly responded to *GA20ox1* overexpression, with altered rosette sizes and longer petioles (Figure 5.2A). Importantly, the response was not always in the same direction. For example, while in the majority of accessions the area of younger leaves were increased, in five accessions (An-1, Ler-0, Blh-1, C24, and WalhaesB4) these leaves were smaller as compared to the corresponding wild-type controls (Figure 5.2B and Extended Data Table 5.2). Overall, eleven accessions had larger rosettes, as measured by 'rosette expressivity' (Ros) (see Methods). Because not all leaves were affected in all lines, three successive leaves with the greatest change in area were examined in more detail (Figure 5.2C and Extended Data Table 5.2) and numerically expressed as 'selective leaf expressivity' (Sel) (see Methods). *GA20ox1* overexpression reduced the size of these three leaves in four accessions, while they were larger in all other accessions. Similarly, pavement cell number and area were affected in a different manner, even when overall leaf sizes showed no changes compared to wild type (Figure 5.2D). In conclusion, *GA20ox1* overexpression causes distinct effects in different accessions with the majority of accessions showing an enhanced leaf and rosette size.

To examine if the accession-specific expressivity could result from differences in sequence of GA20ox1 between the accessions and Col-0, the cDNA and protein sequences were compared (Extended Data Figure 5.5). At DNA level, some sequence differences were found, but, in most of the cases these differences led to synonymous substitutions, therefore not affecting the protein sequence, suggesting that the differences in sequence might not affect the activity of the transgene in the transgenic lines and therefore not influence the expressivity.

Next, we measured GA levels (Figure 5.2E and Extended Data Figure 5.4). As expected, the accumulation of GA20ox1 substrates GA_{53}, GA_{44}, GA_{19} and GA_{24} was reduced, while GA20ox1 products GA_9 and GA_{20}, two bio-active forms, as well as GA_8, a bio-inactive form of GA, were strongly increased in all transgenic lines compared to its wild type control. We noticed that the levels of GA_1 and GA_4, the final bioactive products, were remarkably similar among the lines of each accession, suggesting that levels of these GAs are particularly well buffered within a given accession against different levels of *GA20ox1* overexpression. There was also no correlation between GA levels and expressivity of the growth-related phenotype, indicating that the downstream growth responses differ across accessions.

We also found that 'rosette expressivity' was significantly positively correlated with leaf number, fresh and dry weight of the wild type accessions and significantly negatively correlated with both vascular complexity and density. Similarly, 'selective leaf' expressivity significantly positively correlates with levels of GA_4 and significantly negatively with vascular density of the wild types (Extended Data Figure 5.6 and Extended Data Figure 5.7). In addition, we did not find

a correlation between the geographical origin of an accession and the expressivity of overexpression of *GA20ox1* (data not shown).

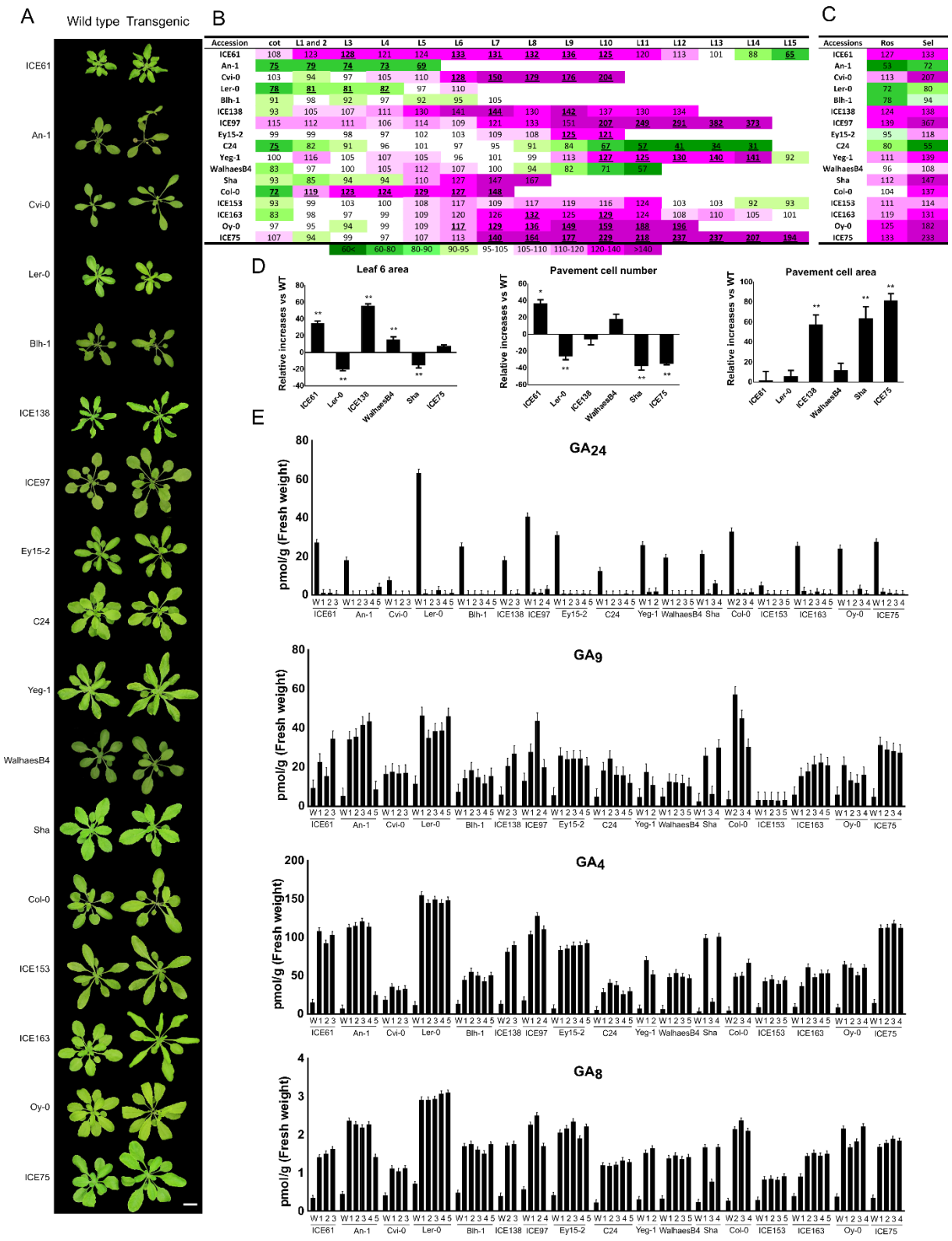


Figure 5.2 Phenotype of *GA20ox1*^{OE} transgenics in 17 Arabidopsis accessions. (A) Image of 25-days-old rosette of representative *GA20ox1* ^{OE} transgenics and their corresponding wild types. Scale bar: 2 cm. (B) Heat map representing, per accession, the average percent difference in each leaf area between *GA20ox1* ^{OE} transgenics and their corresponding wild type. Bold with underline: p-value < 0.05. (C) Heat map

showing the estimated expressivity (Sel: selective leaf, Ros: rosette, see Methods) of *GA20ox1* ^{OE}. (D) Pavement cell number and area in transgenic lines from six different accessions (N=3, * P < 0.05, ** P < 0.01). (E) GA levels in *GA20ox1*^{OE} lines. The normalized values are represented with standard error bars (N=3). W; wild type, 1-5; independent transgenic lines.

Accession-specific transcriptome changes in response to *GA20ox1*^{OE}

We used RNA-seq of ten accessions and their representative transgenic derivatives with variable changes in leaf 6 area to profile differential downstream responses of GA20ox1. Because cell proliferation and/or cell expansion are affected in the transgenic lines (Figure 5.2D) and the transition between cell proliferation and cell expansion is crucial for determining final leaf size (Andriankaja et al., 2012; Gonzalez et al., 2012; Hepworth and Lenhard, 2014), leaves were micro-dissected (size < 0.25 mm^2) at the beginning of this transition, either 12 or 13 DAS depending on the accession (see Methods) and used for RNA-seq.

RNA-seq confirmed overexpression of *GA20ox1* in all transgenic lines (Extended Data Figure 5.8), but this was not predictive of bioactive GA_4 levels As expected, accession-specific properties dominated over the effects of *GA20ox1* overexpression, as deduced from Principal Component Analysis (PCA) (Figure 5.3A). Importantly, similar to what was observed at the morphological level, the direction of the effects of *GA20ox1* overexpression on the transcriptome differed among accessions.

767 genes were differentially expressed (DE) in at least one accession. Overrepresented Gene Ontology (GO) categories were photosynthesis, secondary metabolism, protein, hormone metabolism, regulation of transcription, transport, amino acid metabolism, and sulfur-assimilation pathways (Extended Data Table 5.3, Figure 5.3B, Figure 5.3C and Extended Data Figure 5.9). Genes involved in GA deactivation and degradation (*GIBBERELLIC ACID METHYLTRANSFERASE 2*, *GA2ox1*, and *GA2ox4*) were up-regulated, and GA biosynthetic genes *GA3ox1* and *GA20ox2* were downregulated in many lines, indicative of feedback regulation (Figure 5.3B). Several genes related to other phytohormones, including JA, ABA, brassinosteroid, auxin, ethylene, and cytokinin, were altered in expression, reflecting extensive cross regulation among hormones (Weiss and Ori, 2007). For example, six small auxin up-regulated RNAs (*SAUR*), two ethylene response factor (*ERF*) and a gene encoding a rate limiting enzyme in ABA biosynthesis, 9-cis-epoxycarotenoid dioxygenase (NCED), are differentially expressed in the *GA20ox1* overexpression lines. Genes related to photosynthesis were mostly down-regulated (Figure 5.3C). Since we had analyzed young developing leaves, a possible explanation is that GA promotes growth and delays the onset of differentiation and the establishment of the photosynthetic apparatus by decreasing chlorophyll content in leaf (Cheminant et al., 2011).

Correlation analyses identified genes that were either significantly positively or negatively correlated with expressivity of morphological effects (Figure 5.4). Eighteen were found to be correlated with ‘selective leaf’ and/or ‘rosette’ expressivity. We speculate that these genes might have important roles in determining the influence of *GA20ox1* overexpression in the different accessions.

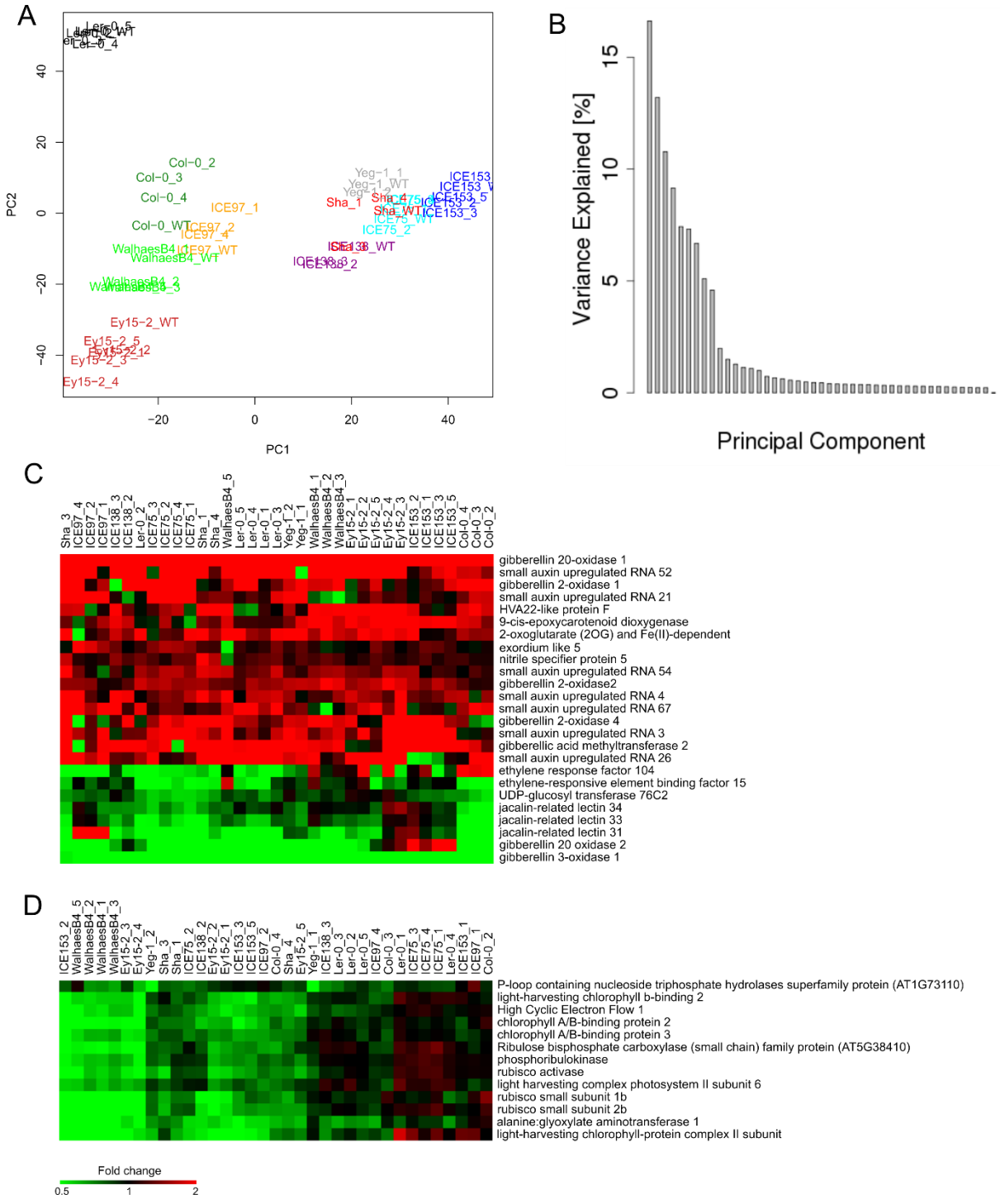


Figure 5.3 Principal Component Analysis (PCA) of transcriptomics data and heat maps representing the fold change of differentially expressed genes in *GA20ox1*^{OE} transgenics. (A) PCA plot represents classifications of transcriptome data wild type and *GA20ox1* ^{OE} transgenics. Each accession is displayed in a different color. W; wild type, 1-5; independent transgenic lines. (B) Plot representing variance explained by PCA axes. Differentially expressed genes involved in hormone metabolism (C) and photosynthesis (D). Red and green colors mean increased and decreased expression in comparison to the wild types, respectively. Only DE genes that show at least 1.5 fold change difference are shown. Hierarchical clustering was done for both genes and samples with Manhattan distance metrics.

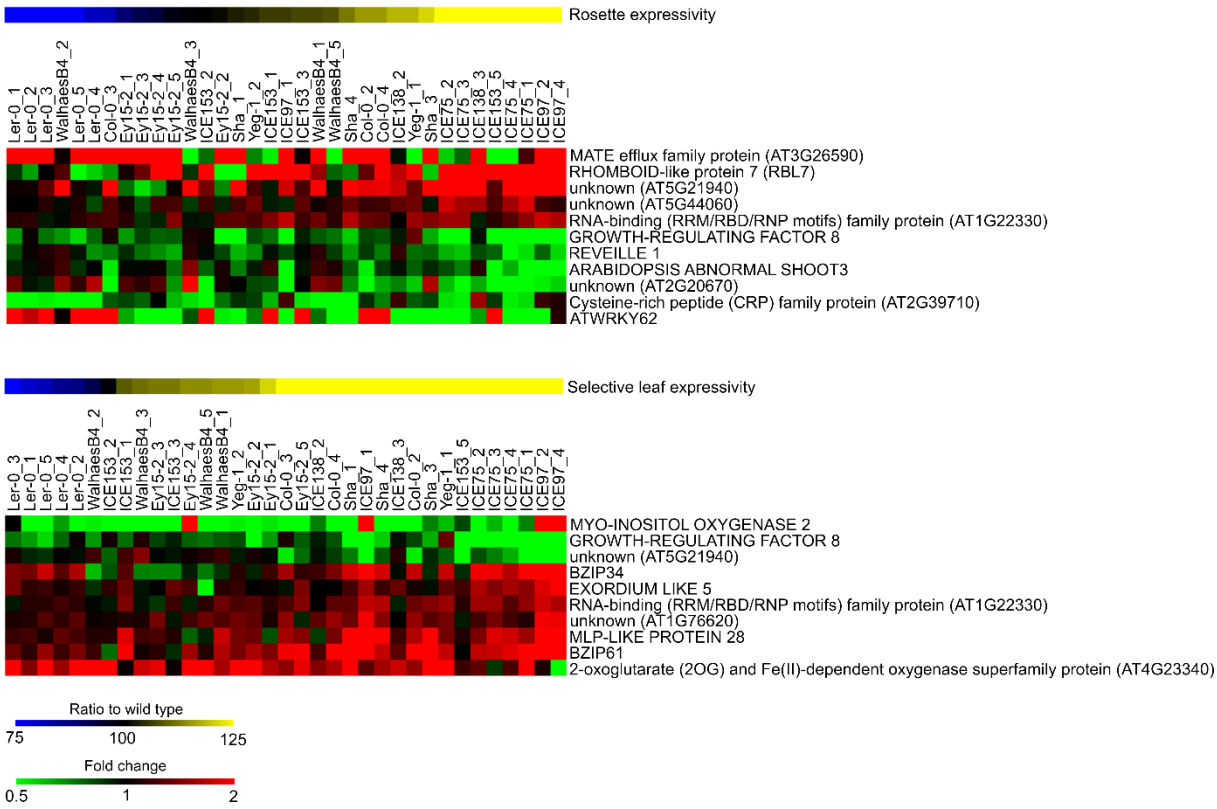


Figure 5.4 Correlation analysis between phenotypic and transcriptome data. Heat maps represent the DE genes correlated with 'rosette' and/or 'selective leaf' expressivities. Upper two heat maps show the expressivity of rosette leaves and DE genes correlated with 'rosette' expressivity. Lower two heat maps represent the expressivity of 'selective leaf' and DE genes correlated with 'selective leaf' expressivity. Red and green colors correspond to increased and decreased expression in comparison to the wild types, respectively. Only DE genes that show at least 1.5 fold change difference are shown. Hierarchical clustering was done for both genes and samples with Manhattan distance metrics. (Correlation coefficient > |0.5|, adj-P value < 0.05)

Discussion

Plant growth is regulated by complex regulatory networks that are determined by the genome and its interaction with ever-changing environment. Such growth regulatory networks are expected to be rather different amongst species and even within a species, which might serve as a key element in adaptation to different environments. It has been demonstrated that mutations or transgenes influencing growth might have different effects in different genetic backgrounds in several model organisms (Dowell et al., 2010; Chandler et al., 2013). Here, we show for 17 different *A. thaliana* accessions, that the ectopic expression of a rate-limiting enzyme for gibberellin biosynthesis has very different effects on growth depending on the accession in which the gene is introduced.

Overexpression of *GA20ox1* in 17 accessions increased the levels of the bioactive forms of GA (GA_1 and GA_4), and depleted GA_{24}, the direct precursor of GA_4, in all accessions, demonstrating that the *GA20ox1* transgene is active in all accessions. A remarkable observation

is that the levels of GA_1 and GA_4 are very similar across multiple transgenic lines of an accession. In other words there appears to be an accession specific maximum accumulation level of GA_1 and GA_4. The reason for this is currently unclear but it is known that bioactive GA forms stimulate the expression and activity of GA catabolism, counteracting the accumulation of the bioactive GAs and converting GA to bio-inactive forms. Furthermore, GA represses the expression of endogenous genes encoding the GA biosynthetic genes *GA20ox* and *GA3ox* (Coles et al., 1999; Nelissen et al., 2012; Ribeiro et al., 2012). Such feedback regulation is also observed in the transcriptome data of the *GA20ox1* overexpression lines in all analysed accessions. Possibly there is an accession specific feedback control in which the GA receptors (GID) and the GA-triggered degradation of DELLA proteins are likely to play a role (Ueguchi-Tanaka et al., 2007; Claeys et al., 2014). The majority of accessions showed upon introduction of the *GA20ox1* transgene a positive effect on the growth of leaves. However the effect quantitatively differs between accessions. Furthermore in some accessions, *GA20ox1* overexpression has even a negative effect on leaf and rosette size. No significant correlation could however be found between the levels of *GA20ox1* overexpression or the levels of various GAs, and the observed effects. Interestingly, the cellular analysis of five *GA20ox1* overexpression lines, all from different accessions, showed different effects at cellular level upon overexpression of *GA20ox1.* Similar genotype-dependent effects on freezing tolerance were found when the cold tolerance genes *CBF1*, *CBF2* and *CBF3* were down regulated in eight different accessions of A. thaliana (Gery et al., 2011).

We observed that biomass of wild type was positively correlated with the growth-promoting effect of *GA20ox1* overexpression on rosette size. Accessions which are, compared to the other accessions, larger show a more pronounced response to *GA20ox1* overexpression than the accessions with smaller rosettes. We hypothesize that in large accessions the growth regulatory network is less constrained and more prone to the effect of positive growth regulators, whereas in small accessions, which have a more restrictive growth network, it would be more difficult to make larger plants. In addition, it seems that there is more room for physical expansion in larger accessions since vascular density and complexity in wild type showed negative correlation with expressivities. For 'rosette' expressivity no direct correlation with GA levels was found, whereas the GA20ox1 effect on selected leaves positively correlate with GA_4 levels. A possible reason for this observation is that rosette size is a complex trait determined by many different factors, amongst which leaf number and size, speed of development while leaf size is a simpler trait, as it does not integrate different individual organs.

How can we explain the natural variation in the effect of *GA20ox1* overexpression based on our finding of almost no strong correlation between its transcript level, active GA quantity and phenotypic effects? Many steps exist from the expression of *GA20ox1* to its actual effect on growth and differences in signal transduction along the GA pathway, depending on the genetic background, could therefore be the reason for the observed variability. First, translatome analysis after treatment with bioactive GA revealed that differential mRNA translation, possibly varying between the different accessions, is important for the control of feedback regulation of

GA-related genes (Ribeiro et al., 2012). Second, at protein level, the amount of the GA-receptor (GID) and DELLAs, negative regulators of GA signalling, their affinity and efficiency of interaction to form the regulatory module GA-GID-DELLA might be different in the different accessions and therefore affect the differential response existing between the accessions. Distinct interactions with other growth regulatory elements could also explain the variation observed. It has been shown, in Col-0, that overexpression of *GA20ox1* in binary combination with an altered expression of growth promoting genes leads to different size phenotypes in function of the gene combination (Vanhaeren et al., 2014). It is therefore possible, that differences in expression of growth regulatory genes in the natural variants, triggering different cellular characteristics in the wild type plants, influence differently the effect of *GA20ox1* overexpression. In addition, we identified 18 genes of which the expression levels are correlated with the phenotypic expressivity of *GA20ox1* in all analysed accessions. Three genes were found to be correlated with both 'rosette' and 'selective leaf' expressivities. A few genes out of 18 genes have previously been associated with plant growth, one of three overlapped gene, *GRF8* is one of the nine members of the *GROWTH REGULATING FACTOR* gene family with major role in regulating cell proliferation and/or cell expansion during plant development (Kim et al., 2003; Vercruyssen et al., 2014). Two auxin related genes were found to correlate with 'rosette' expressivity: *ARABIDOPSIS ABNORMAL SHOOT3* (Li et al., 2014) and *REVEILLE1* (Rawat et al., 2009). Members of the *EXORDIUM* gene family correlating with 'selective leaf' expressivity were previously found to have a role in cell expansion and brassinosteroid signalling (Schröder et al., 2009). Further work is required to determine whether this subset of genes have a functional role in determining the accession specific responses to elevated GA levels. In order to provide further insight in the mechanism that is behind the accession specific effect of GA perturbation, screening for modifier genes that suppress the response to GA perturbation in transgenic lines of a specific accession could be performed. Furthermore, detailed analysis of the GA signalling pathway in the different accessions is likely to shed light on how GA affects growth to a very different extent in different Arabidopsis accessions.

Methods

Plant Material and Growth Conditions

Seventeen Arabidopsis accessions were selected to cover most common genetic variants of *Arabidopsis thaliana* (Table 4.1) and used to generate *GA20ox1* overexpressing lines. cDNA of the full *GA20ox1* coding region from Col-0 was cloned in the fluorescence-accumulating seed technology (FAST) vectors (Shimada et al., 2010) and introduced into the 17 accessions following the floral dip protocol (Clough and Bent, 1998). Dried transgenic T_1 seeds were selected based on fluorescence signal in the seed coat and sown on soil for seed production. T_2 transgenic seeds were harvested and selection of 5 independent single locus insertion lines (75% of seeds

fluorescent) was done. Seeds were sown on soil for seed production and expression of the transgene was verified by RT-qPCR. From these lines, at least 2 and maximum 5 independent T_3 homozygote lines for each accession were selected for further experiments. All plants were grown on the MIRGIS platform in soil under a 16-hours-day/8-hours-night regime at 21°C in a growth chamber. The daily images of plants obtained from MIRGIS are analyzed in Chapter 8.

Phenotypic analysis

All phenotypic measurements, including shoot biomass (fresh and dry weight), were obtained from 25-day old plants. Leaf series were made by dissecting individual leaves (from cotyledon to the younger rosette leaf) and mounting them on 1 % agar plate, and the leaf area was measured with the ImageJ software (http://rsb.info.nih.gov/ij/). The leaf series for the 17 wild type accessions were done in three biological repeats, while the leaf series for the transgenics were in one biological repeat by growing the transgenics together with their corresponding wild type. Measurements of venation patterns were done as previously described (Dhondt et al., 2012) from leaf 1 and 2. Ploidy levels of leaf 1 and 2 were measured and endoreduplication index was calculated as previously described (Claeys et al., 2012). For the heat map of leaf-size related parameters (total rosette area, fresh weight and dry weight of the shoot, total number and area of leaves, pavement cell number and area, cell circularity, endoreduplication index, stomatal density, stomatal index, and vascular complexity and density of leaf 1 and 2) in 17 accessions, the measured value for a parameter in each accession was divided by the average of this parameter for all accessions.

For the leaf series data, leaf area was log transformed to stabilize the variance. The mean model consisted of the main effects of overexpression of *GA20ox1* and leaf position and their interaction term. Due to the unbalanced and complex nature of the data, the Kenward-Rogers approximation for computing the denominator degrees of freedom for the tests of fixed effects was used. An autoregressive structure was used to model the correlations between measurements done on the leaves originating from the same plant. The main interest was in the effect of the gene on leaf area for each leaf separately. Simple tests of effects were performed at each leaf between the transgenics and the wild type. Difference estimates were represented as % to the least-square means estimate of the wild type and leaf. Separate models were run for each accession as they were grown in separate experiments. For each experiment, the data was truncated so that there were at least 2 observations for each leaf of both the transgenic and the wild type. The analysis was performed with the mixed and plm procedure of SAS (Version 9.4 of the SAS System for windows 7 64bit. Copyright © 2002-2012 SAS Institute Inc. Cary, NC, USA (www.sas.com). Residual diagnostics were carefully examined.

The expressivity of 'selective leaf' was determined by averaging of the ratio of three successive leaves showing the highest change in area. Same leaves were taken for each accession to estimate 'selective leaf' expressivity. 'Rosette' expressivity was estimated as a ratio

of wild type rosette area to a transgenic line rosette area. In case of 'rosette' expressivity per accession the mean of 'rosette' expressivity per transgenic of an accession has been taken.

Cellular analysis was done as previously described (Andriankaja et al., 2012) in three repeats.

For transgenic lines, significant differences were determined with a two-way ANOVA test with genotype and repeat as main factors in R software (v 3.0.1) (R Core Team, 2015). Differences between the wild type and corresponding transgenic line were estimated and declared significant when adj-p value < 0.05 with Tukey's method.

Hormone Analysis

For hormone measurements, we chose Arabidopsis development stage 1.03 (Boyes et al., 2001) (12 DAS for Cvi-0 and 11 DAS for the other accessions) since gibberellins are produced in young developmental stages of Arabidopsis. The shoot of seedlings grown in soil was harvested in the middle of the day from three biological independent experiments and frozen in liquid nitrogen. The phytohormones GA (GA_4, GA_8, GA_9, GA_{19}, GA_{24}, GA_{44}, and GA_{53}), IAA (IAA, IAAsp, IAIle + IALeu, and IAPhe), ABA, SA, cytokinin (tZ, tZR, tZRPs, cZ, cZR, cZRPs, DZ, DZR, DZRPs, iP, iPR, iPRPs, tZ7G, tZ9G, tZOG, tZROG, cZROG, tZRPsOG, DZ9G, iP7G, and iP9G), and JA were measured as described previously (Kojima et al., 2009; Shinozaki et al., 2015). The hormone data were modelled with a linear model with accession as main factor and experiment as fixed block factor due to small number of samples (three repeats). The model was fitted with the lm function from the R software (v 3.0.1) (R Core Team, 2015). Least-squares means and standard errors were calculated with the lsmeans function of the lsmeans library (v. 2.10) (Lenth and Hervé, 2014) from the R software (v 3.0.1) (R Core Team, 2015). These estimates were used in Pearson correlation analyses.

RNA Extraction

Total RNA was extracted from the shoot of 12-day old seedlings of T_2 transgenic lines and the corresponding wild type plants using trizol and the expression of the transgene was analysed by Quantitative reverse transcription-PCR (RT-qPCR). RT-qPCR was performed as previously described (Claeys et al., 2012).

For RNA sequencing (RNA-seq) analysis, seedlings with one biological repeat of wild-type plants and GA20OX overexpressing lines (at 12 DAS for Col-0 and Ey15-2 and at 13 DAS for WalhaesB4, ICE97, ICE138, ICE75, Ler-0, Yeg-1, Sha, and ICE153) were harvested in RNA ice-later solution (AM7030; Ambion) and incubated at -20°C for at least a week. Leaf 6 was micro-dissected on a cold plate with dry ice under a stereomicroscope, and frozen in liquid nitrogen. RNA was extracted with Trizol (Invitrogen) according to the manufacturer's instructions and the RNeasy kit (Qiagen) with on-column DNase (Qiagen) digestion. RNA was quantified and the quality was checked with a 2100 Bioanalyzer (Agilent).

RNA sequencing Analysis

Library preparation was done using the TruSeq RNA Sample Preparation Kit v2 (Illumina). In brief, polyA containing mRNA molecules are reverse transcribed, double-stranded cDNA is generated and adapters are ligated. After quality control using 2100 Bioanalyzer (Agilent), clusters are generated through amplification using the TruSeq PE Cluster Kit v3-cBot-HS kit (Illumina) followed by sequencing on a Illumina HiSeq2000 with the TruSeq SBS Kit v3-HS (Illumina). Sequencing was performed in Paired-End mode with a read length of 50 bp. The quality of the raw data was verified with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/, version 0.9.1). Next, quality filtering was performed using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/, version 0.0.13): reads where globally filtered in which for at least 75% of the reads the quality exceeds Q20 and 3' trimming was performed to remove bases with a quality below Q10. Re-pairing was performed using a custom perl script. Reads were subsequently mapped to the Arabidopsis reference genome (TAIR10) using GSNAP (Wu and Nacu, 2010) (version 2011-12-28) allowing maximally 2 mismatches. The concordantly paired reads that uniquely map to the genome were used for quantification on the gene level with htseq-count from the HTSeq.py python package (Anders et al., 2015). The analysis was implemented as a workflow in Galaxy (Goecks et al., 2010).

For the visualization of RNA-Seq expression data and correlation analysis, count data was normalized following the normalization pipeline with the trimmed mean of M-values (TMM) algorithm as implemented in the edgeR library from the R software (v.3.0.1) (R Core Team, 2015). Weakly expressed genes were previously filtered out by removing genes that have less than 5 samples with expression lower than 10 counts per million. The normalized count data was then transformed with inverse hyperbolic sine function "asinh" in R software (v.3.0.1) (R Core Team, 2015).

The PCA plot on transformed count data was done in R using 'pca' function.

Sequence extraction and alignment

The different read libraries were quality checked using FastQC v0.9.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapters were trimmed using cutadapt in python v3.1.1 with the options: --adapter 'GATCGGAAGAGCACACGTCTGAACTCCAGTCAC', --overlap 10 and --minimum-length 35 (Martin, 2011). Quality filtering was performed using Fastx v0.0.13 with the options: -Q 33, -q 10 and -p 75 (http://hannonlab.cshl.edu/fastx_toolkit/). Reads were also trimmed using Fastx v0.0.13 with the options: -Q 33, -t 20 and -l 35. Forward and reverse reads were subsequently collapsed into a single file. After preprocessing, the different read libraries were mapped to the TAIR10 genome using gsnap v2013-02-05 with the options --trim-mismatch-score -3, -k 15, -A sam, -B 4, -n 50, -w 15000, -a off, --quality-protocol sanger, --pairmax-rna 15000, -N 1 and -m 5 (Wu and Nacu, 2010).

Only uniquely mapping reads were further considered. Next, sorting and deduplication of the read libraries was performed using picard v1.129 (http://broadinstitute.github.io/picard/). GATK v3.3.0 was used for variant calling (Van der Auwera et al., 2013). Analysis was based on recommendations in 'Best practices for RNA-seq' (https://www.broadinstitute.org/gatk/guide/best-practices?bpm=RNAseq). Before variant calling was performed, the different libraries were preprocessed using the tools splitnCigar, haplotypecaller, realignertargetcreator, indelrealigner, baserecalibrator and printreads. In the haplotypecaller step only high quality scores were considered by setting a quality of 50. Next, a multi-sample variant calling was performed using haplotypecaller. In this step, all samples are analysed together. Variants were filtered using VariantFiltration with the options -window 35, -cluster 3, -filterName FS, -filter "FS > 30.0", -filterName QD and -filter "QD < 2.0". The resulting variants file was split by sample using bcftools (http://github.com/samtools/bcftools). Sequences were extracted for the genes (AT4G14713, AT4G14720, AT1G19270, AT2G01570 and AT4G25420) using the alternative alleles for each sample using the GATK tool Fasta Alternate Reference Maker (Van der Auwera et al., 2013) and based on the CDS coordinates (based on the structural annotation of TAIR10). The reverse complement was generated for genes located at the negative strand and subsequently protein sequences were extracted using custom scripts.

To align extracted sequence, CLC main Workbench 6.0 was used (CLC bio, a QIAGEN Company; http://www.clcbio.com/).

Differential Expression Analysis

Differential expression analyses of RNA-Seq data were conducted with the 'EdgeR' library (v.3.4.2) of the Bioconductor software from the R software (v.3.0.1) (R Core Team, 2015). Filtering and normalization was performed as previously described. In this analysis we consider transgenic lines of a particular accession as repeats of a single line, otherwise we would not be able to run statistical tests as we have a single repeat per line. Two statistical tests were performed; the first test for general, mean differential expression between wild types of accessions and transgenic lines of these accessions. The second test is for genes differentially expressed between wild type of an accession and a transgenic line of that accession in at least one accession. The tests were run using the glmLRT function with a contrast adequate to the test. Next, p-values from both tests were corrected for multiple testing with FDR using the qvalue function from 'qvalue' package (v.1.36.0) (Storey et al., 2015) for Bioconductor. This results in two q-values (corrected p-values from both tests) for each gene. The lowest value was assigned as the final q-value for a gene. For further analysis genes were selected based on a FDR value lower than 0.05 and/or fold change threshold between transgenic lines and wild type. The filter on the fold change requires fold change higher than 1.5 for each transgenic line of an accession in at least one accession.

Enrichment analysis was done in mapman (Ramšak et al., 2014) (http://mapman.mpimp-golm.mpg.de/pageman/) with DE genes filtered for FDR lower than 0.05.

Heat maps are generated in Mev (v 4.9) (Howe et al., 2011) for DE gene filtered for FDR lower than 0.05 and 1.5 fold change threshold between transgenic lines and wild type. Hierarchical clustering was done for both genes and samples with Manhattan distance metrics in Mev (v 4.9) (Howe et al., 2011).

Correlation Analysis

Pearson correlations coefficients were calculated with corr.test function in R. The adjusted P values of correlations were calculated with a permutation test. For each analyses phenotype-phenotype, phenotype-hormone, hormone-hormone, expressivity-RNA-seq fold change we run 1000 times permutations, where we permute either one phenotype or metabolite or expressivity. Permutation analysis is conducted for every phenotype/hormone/expressivity independently. The adjusted P values are calculated as a proportion of correlation coefficients higher correlated than a tested correlation (r) to the number of permutations (n); with a formula (r+1)/(n+1) (North et al., 2002). The significant correlations, FDR<0.05, were visualized in Cytoscape (Cline et al., 2007).

Extended Data

All extended data is listed below. Extended data can be found at the end of this chapter.

Extended Data Figure 5.1 Correlation between the shoot-related phenotypic measurements of the 17 Arabidopsis accessions.

Extended Data Figure 5.2 Correlation between four of the major bioactive hormones (ABA, cytokinins, JA, and ABA) in 17 Arabidopsis accessions.

Extended Data Figure 5.3 Correlation between the leaf size-related parameters and hormones in the 17 Arabidopsis accessions.

Extended Data Figure 5.4 GA levels in *GA20ox1* ^{OE} lines from the 17 Arabidopsis accessions.

Extended Data Figure 5.5 Sequence alignments of the endogenous cDNA and protein sequence of GA20ox1 in 17 *Arabidopsis* accessions.

Extended Data Figure 5.6 Correlation analysis of phenotypic data.

Extended Data Figure 5.7 Scatter plots showing the correlation between the leaf size-related parameters in 17 Arabidopsis accessions and the expressivities.

Extended Data Figure 5.8 *GA20ox1* expression level in the transgenic lines from 10 Arabidopsis accessions.

Extended Data Figure 5.9 Heat maps representing the fold change of DE genes in *GA20ox1* ^{OE} lines.

Extended Data Table 5.1 Correlation between levels of different hormones in the 17 Arabidopsis accessions.

Extended Data Table 5.2 Phenotype of *GA20ox1*^{OE} transgenics in 17 Arabidopsis accessions per independent transgenic line.

Extended Data Table 5.3 Over-represented MapMan categories for *GA20ox1* DE genes.

References

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq – a Python framework to work with high-throughput sequencing data. Bioinformatics **31:** 166-169.

Andriankaja, M., Dhondt, S., De Bodt, S., Vanhaeren, H., Coppens, F., De Milde, L., Mühlenbock, P., Skirycz, A., Gonzalez, N., Beemster, G.T.S., and Inzé, D. (2012). Exit from proliferation during leaf development in *Arabidopsis thaliana*: a not-so-gradual process. Dev. Cell **22:** 64-78.

Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R., and Görlach, J. (2001). Growth stage-based phenotypic analysis of Arabidopsis: a model for high throughput functional genomics in plants. Plant Cell **13:** 1499-1510.

Chandler, C.H., Chari, S., and Dworkin, I. (2013). Does your gene need a background check? How genetic background impacts the analysis of mutations, genes, and evolution. Trends Genet. **29:** 358-366.

Chari, S., and Dworkin, I. (2013). The conditional nature of genetic interactions: the consequences of wild-type backgrounds on mutational interactions in a genome-wide modifier screen. PLoS Genet. **9:** e1003661.

Cheminant, S., Wild, M., Bouvier, F., Pelletier, S., Renou, J.-P., Erhardt, M., Hayes, S., Terry, M.J., Genschik, P., and Achard, P. (2011). DELLAs regulate chlorophyll and carotenoid biosynthesis to prevent photooxidative damage during seedling deetiolation in *Arabidopsis*. Plant Cell **23:** 1849-1860.

Claeys, H., De Bodt, S., and Inzé, D. (2014). Gibberellins and DELLAs: central nodes in growth regulatory networks. Trends Plant Sci. **19:** 231-239.

Claeys, H., Skirycz, A., Maleux, K., and Inzé, D. (2012). DELLA signaling mediates stress-induced cell differentiation in Arabidopsis leaves through modulation of anaphase-promoting complex/cyclosome activity. Plant Physiol. **159:** 739-747.

Cline, M.S., et al. (2007). Integration of biological networks and gene expression data using Cytoscape. Nat. Protoc. **2:** 2366-2382.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:** 735-743.

Coles, J.P., Phillips, A.L., Croker, S.J., García-Lepe, R., Lewis, M.J., and Hedden, P. (1999). Modification of gibberellin production and plant development in *Arabidopsis* by sense and antisense expression of gibberellin 20-oxidase genes. Plant J. **17:** 547-556.

De Bruyne, L., Höfte, M., and De Vleesschauwer, D. (2014). Connecting growth and defense: the emerging roles of brassinosteroids and gibberellins in plant innate immunity. Mol. Plant **7:** 943-959.

Dhondt, S., Van Haerenborgh, D., Van Cauwenbergh, C., Merks, R.M.H., Philips, W., Beemster, G.T.S., and Inzé, D. (2012). Quantitative analysis of venation patterns of Arabidopsis leaves by supervised image analysis. Plant J. **69:** 553-563.

Dowell, R.D., et al. (2010). Genotype to phenotype: a complex problem. Science **328:** 469.

Félix, M.-A. (2007). Cryptic quantitative evolution of the vulva intercellular signaling network in *Caenorhabditis*. Curr. Biol. **17:** 103-114.

Félix, M.-A., and Wagner, A. (2008). Robustness and evolution: concepts, insights and challenges from a developmental model system. Heredity **100:** 132-140.

Gery, C., Zuther, E., Schulz, E., Legoupi, J., Chauveau, A., McKhann, H., Hincha, D.K., and Téoulé, E. (2011). Natural variation in the freezing tolerance of *Arabidopsis thaliana*: effects of RNAi-induced CBF depletion and QTL localisation vary among accessions. Plant Sci. **180:** 12-23.

Goecks, J., Nekrutenko, A., Taylor, J., and The Galaxy Team. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol. **11:** R86.

Gonzalez, N., Vanhaeren, H., and Inzé, D. (2012). Leaf size control: complex coordination of cell division and expansion. Trends Plant Sci. **17:** 332-340.

Gonzalez, N., et al. (2010). Increased leaf size: different means to an end. Plant Physiol. **153:** 1261-1279.

Hedden, P. (2003). The genes of the Green Revolution. Trends Genet. **19:** 5-9.

Hedden, P., and Thomas, S.G. (2012). Gibberellin biosynthesis and its regulation. Biochem. J. **444:** 11-25.

Hepworth, J., and Lenhard, M. (2014). Regulation of plant lateral-organ growth by modulating cell number and size. Curr. Opin. Plant Biol. **17:** 36-42.

Howe, E.A., Sinha, R., Schlauch, D., and Quackenbush, J. (2011). RNA-Seq analysis in MeV. Bioinformatics **27:** 3209-3210.

Huang, S., Raman, A.S., Ream, J.E., Fujiwara, H., Cerny, R.E., and Brown, S.M. (1998). Overexpression of 20-oxidase confers a gibberellin-overproduction phenotype in Arabidopsis. Plant Physiol. **118:** 773-781.

Kim, J.H., Choi, D., and Kende, H. (2003). The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in *Arabidopsis*. Plant J. **36:** 94-104.

Kojima, M., Kamada-Nobusada, T., Komatsu, H., Takei, K., Kuroha, T., Mizutani, M., Ashikari, M., Ueguchi-Tanaka, M., Matsuoka, M., Suzuki, K., and Sakakibara, H. (2009). Highly sensitive and high-throughput analysis of plant hormones using MS-probe modification and liquid chromatography-tandem mass spectrometry: an application for hormone profiling in *Oryza sativa*. Plant Cell Physiol. **50:** 1201-1214.

Lenth, R.V., and Hervé, M. (2014). lsmeans: Least-Squares Means. R package version 2.13 (https://cran.r-project.org/web/packages/lsmeans/index.html).

Li, R., et al. (2014). *ADP1* affects plant architecture by regulating local auxin biosynthesis. PLoS Genet. **10:** e1003954.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal **17:** 10-12.

Nelissen, H., Rymen, B., Jikumaru, Y., Demuynck, K., Van Lijsebettens, M., Kamiya, Y., Inzé, D., and Beemster, G.T.S. (2012). A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. Curr. Biol. **22:** 1183-1187.

North, B.V., Curtis, D., and Sham, P.C. (2002). A note on the calculation of empirical *P* values from Monte Carlo procedures. Am. J. Hum. Genet. **71:** 439-441.

Paaby, A.B., White, A.G., Riccardi, D.D., Gunsalus, K.C., Piano, F., and Rockman, M.V. (2015). Wild worm embryogenesis harbors ubiquitous polygenic modifier variation. eLife **4:** e09178.

R Core Team. (2015). R: a language and environment for statistical computing. Vienna, Austria (http://www.R-project.org/).

Ramšak, Ž., Baebler, Š., Rotter, A., Korbar, M., Mozetič, I., Usadel, B., and Gruden, K. (2014). GoMapMan: integration, consolidation and visualization of plant gene annotations within the MapMan ontology. Nucleic Acids Res. **42:** D1167-1175.

Rawat, R., Schwartz, J., Jones, M.A., Sairanen, I., Cheng, Y., Andersson, C.R., Zhao, Y., Ljung, K., and Harmer, S.L. (2009). REVEILLE1, a Myb-like transcription factor, integrates the circadian clock and auxin pathways. Proc. Natl. Acad. Sci. USA **106:** 16883-16888.

Ribeiro, D.M., Araújo, W.L., Fernie, A.R., Schippers, J.H., and Mueller-Roeber, B. (2012). Translatome and metabolome effects triggered by gibberellins during rosette growth in *Arabidopsis*. J. Exp. Bot. **63:** 2769-2786.

Schilsky, R.L. (2010). Personalized medicine in oncology: the future is now. Nat. Rev. Drug Discov. **9:** 363-366.

Schröder, F., Lisso, J., Lange, P., and Müssig, C. (2009). The extracellular EXO protein mediates cell expansion in Arabidopsis leaves. BMC Plant Biol. **9:** 20.

Schwechheimer, C., and Willige, B.C. (2009). Shedding light on gibberellic acid signalling. Curr. Opin. Plant Biol. **12:** 57-62.

Shimada, T.L., Shimada, T., and Hara-Nishimura, I. (2010). A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. Plant J. **61:** 519-528.

Shinozaki, Y., et al. (2015). Ethylene suppresses tomato (*Solanum lycopersicum*) fruit set through modification of gibberellin metabolism. Plant J. **83:** 237-251.

Storey, J.D., Bass, A.J., Dabney, A., and Robinson, D. (2015). qvalue: Q-value estimation for false discovery rate control. R package version 2.2.0 (http://www.bioconductor.org/packages/release/bioc/html/qvalue.html).

Ueguchi-Tanaka, M., Nakajima, M., Motoyuki, A., and Matsuoka, M. (2007). Gibberellin receptor and its role in gibberellin signaling in plants. Annu. Rev. Plant Biol. **58:** 183-198.

Van der Auwera, G.A., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics **11:** 11 10 11-11 10 33.

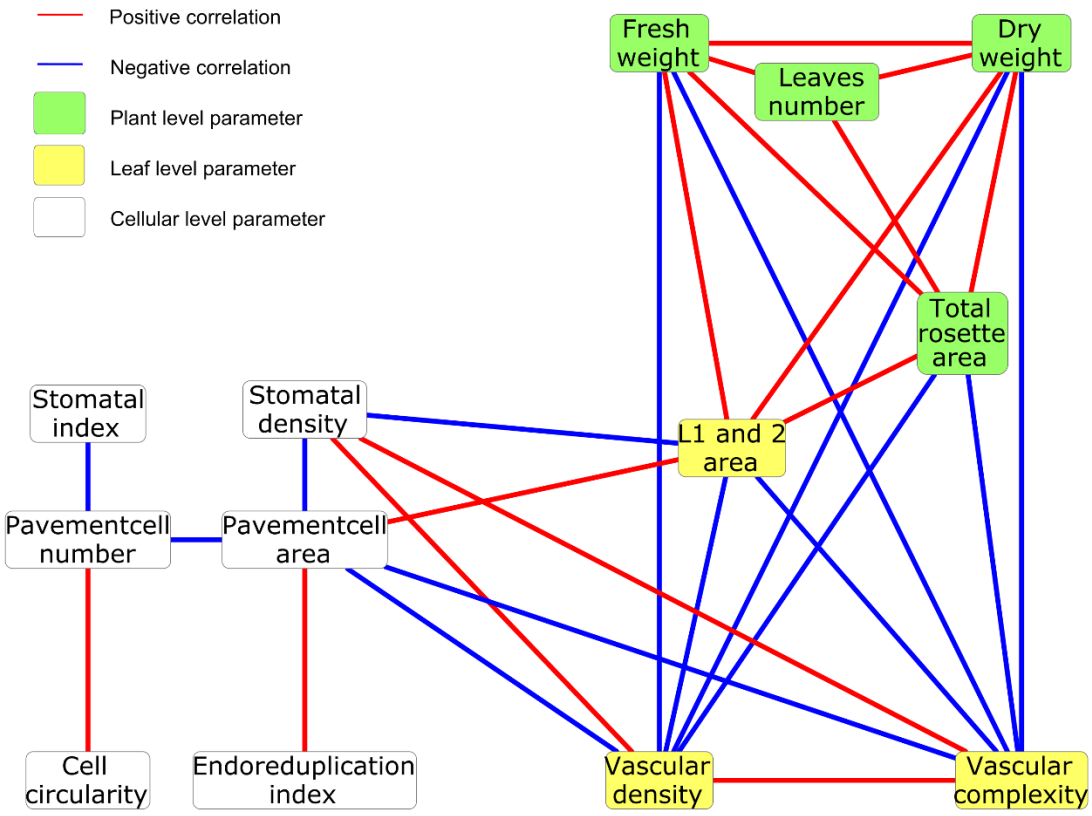
Vanhaeren, H., Gonzalez, N., Coppens, F., De Milde, L., Van Daele, T., Vermeersch, M., Eloy, N.B., Storme, V., and Inzé, D. (2014). Combining growth-promoting genes leads to positive epistasis in *Arabidopsis thaliana*. eLife **3:** e02252.

Vercruyssen, L., et al. (2014). ANGUSTIFOLIA3 binds to SWI/SNF chromatin remodeling complexes to regulate transcription during *Arabidopsis* leaf development. Plant Cell **26:** 210-229.

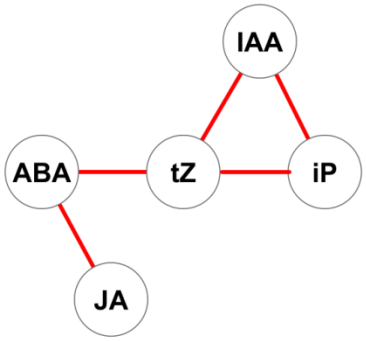
Vu, V., Verster, A.J., Schertzberg, M., Chuluunbaatar, T., Spensley, M., Pajkic, D., Hart, G.T., Moffat, J., and Fraser, A.G. (2015). Natural variation in gene expression modulates the severity of mutant phenotypes. Cell **162:** 391-402.

Weiss, D., and Ori, N. (2007). Mechanisms of cross talk between gibberellin and other hormones. Plant Physiol. **144:** 1240-1246.

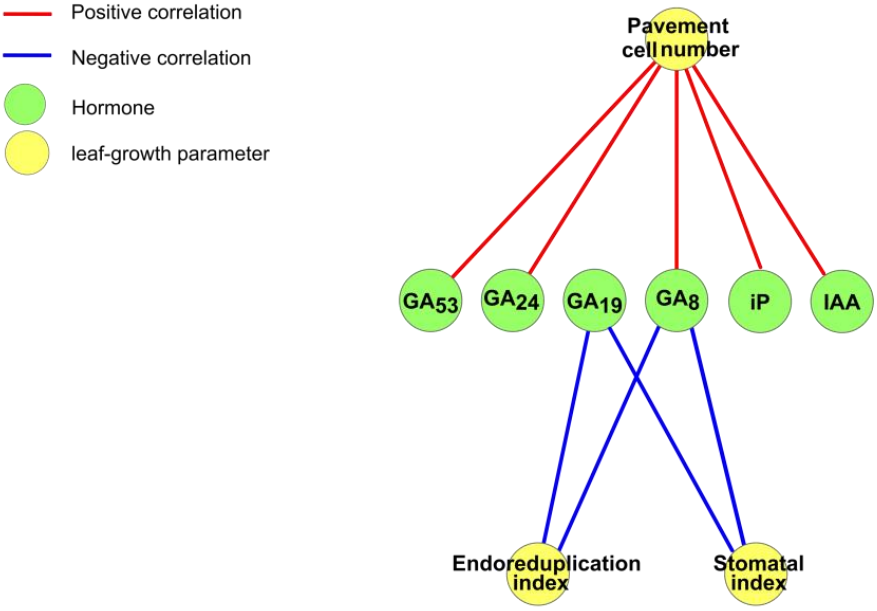
Wu, T.D., and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics **26:** 873-881.



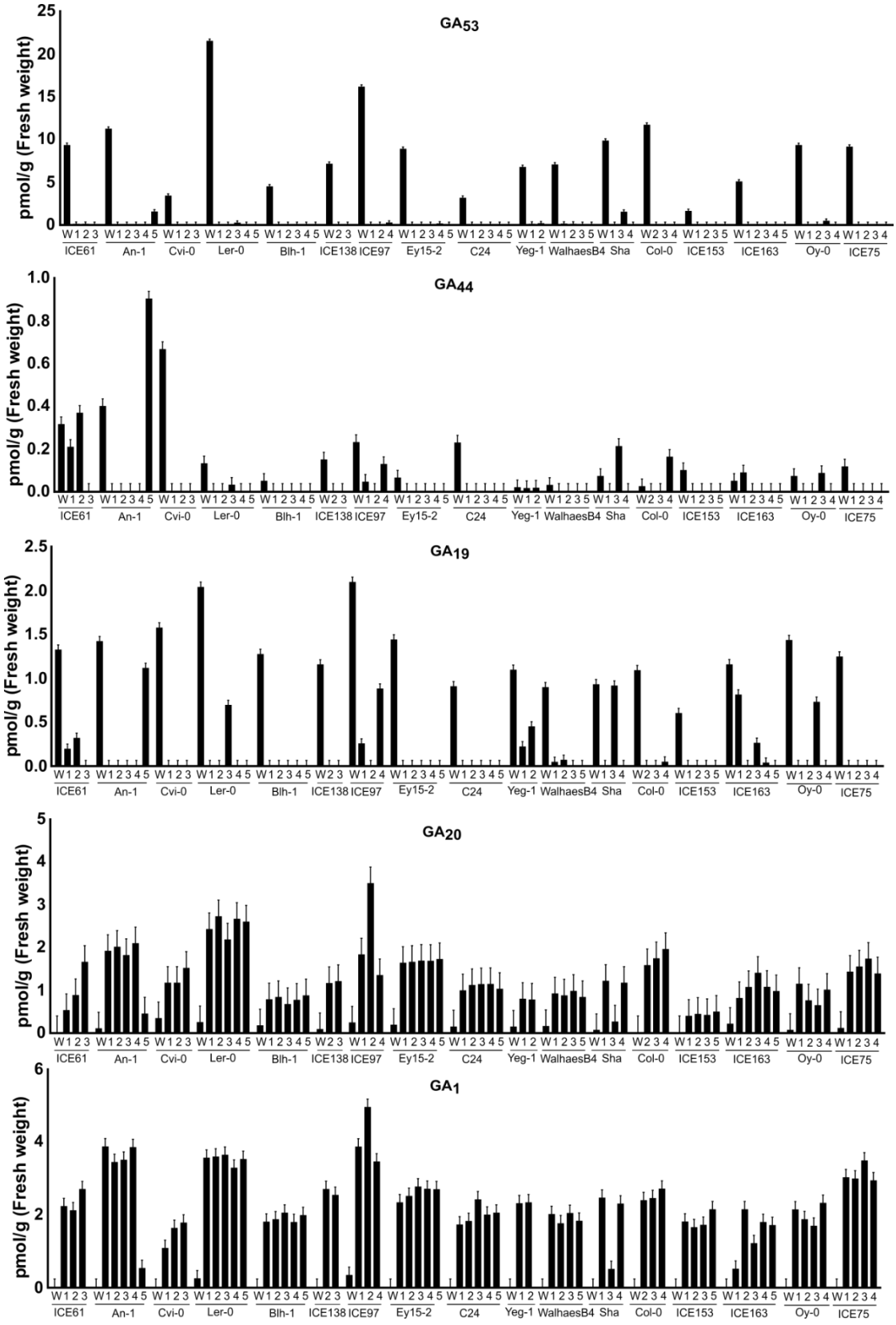
Extended Figure 5.1 Correlation between the shoot-related phenotypic measurements of the 17 Arabidopsis accessions. The parameters measured are fresh and dry weight; total rosette area; total number and area of leaves; pavement cell number, area, and circularity; endoreduplication index; stomatal density and index; vascular complexity and density of the first leaf pair. The green, yellow, and white nodes represent the parameters at plant, leaf, and cellular level, respectively. The cellular level parameters were measured from leaf 1 and 2. The red and blue edges show positive (correlation coefficient > 0.6) and negative correlation (correlation coefficient < -0.6) between parameters, respectively (adj-P value < 0.05).



Extended Figure 5.2 Correlation between four of the major bio-active hormones (ABA, cytokinins, JA, and ABA) in 17 Arabidopsis accessions. The edges indicate positive correlation (correlation coefficient > 0.6) between the hormones (adj-P value < 0.05).

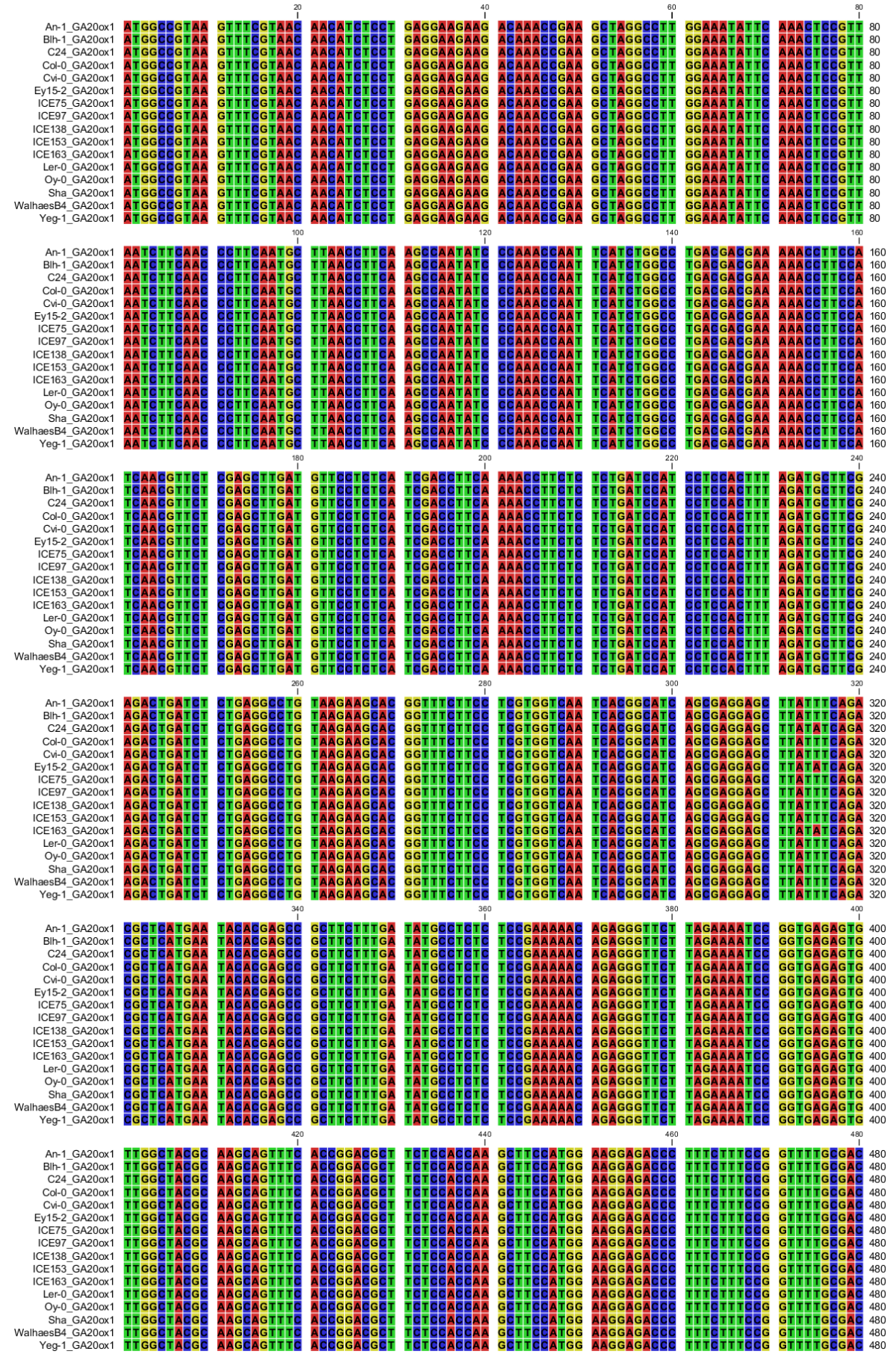


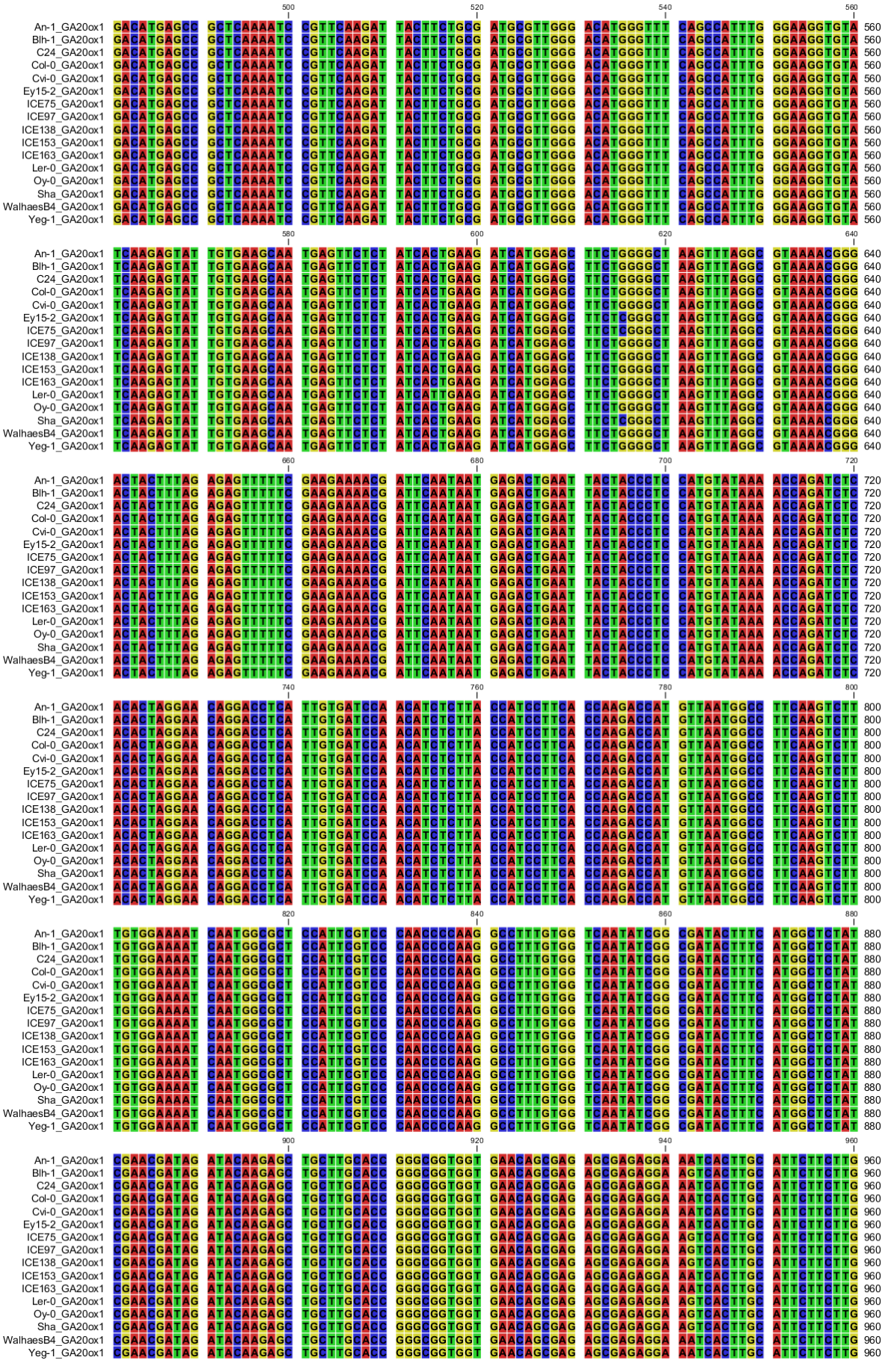
Extended Figure 5.3 Correlation between the leaf size-related parameters and hormones in the 17 Arabidopsis accessions. The red and blue edges show positive (correlation coefficient > 0.6) and negative correlation (correlation coefficient < -0.6) between parameters, respectively (adj-P value < 0.05).



Extended Figure 5.4 GA levels in *GA20ox1*^{OE} lines from the 17 Arabidopsis accessions. GA_{53}, GA_{44}, GA_{19}, GA_{20} and GA_1 were measured from 12–day-old seedlings grown in soil. The normalized values are represented with standard error bars (N=3). W; wild type, 1-5; independent transgenic lines.

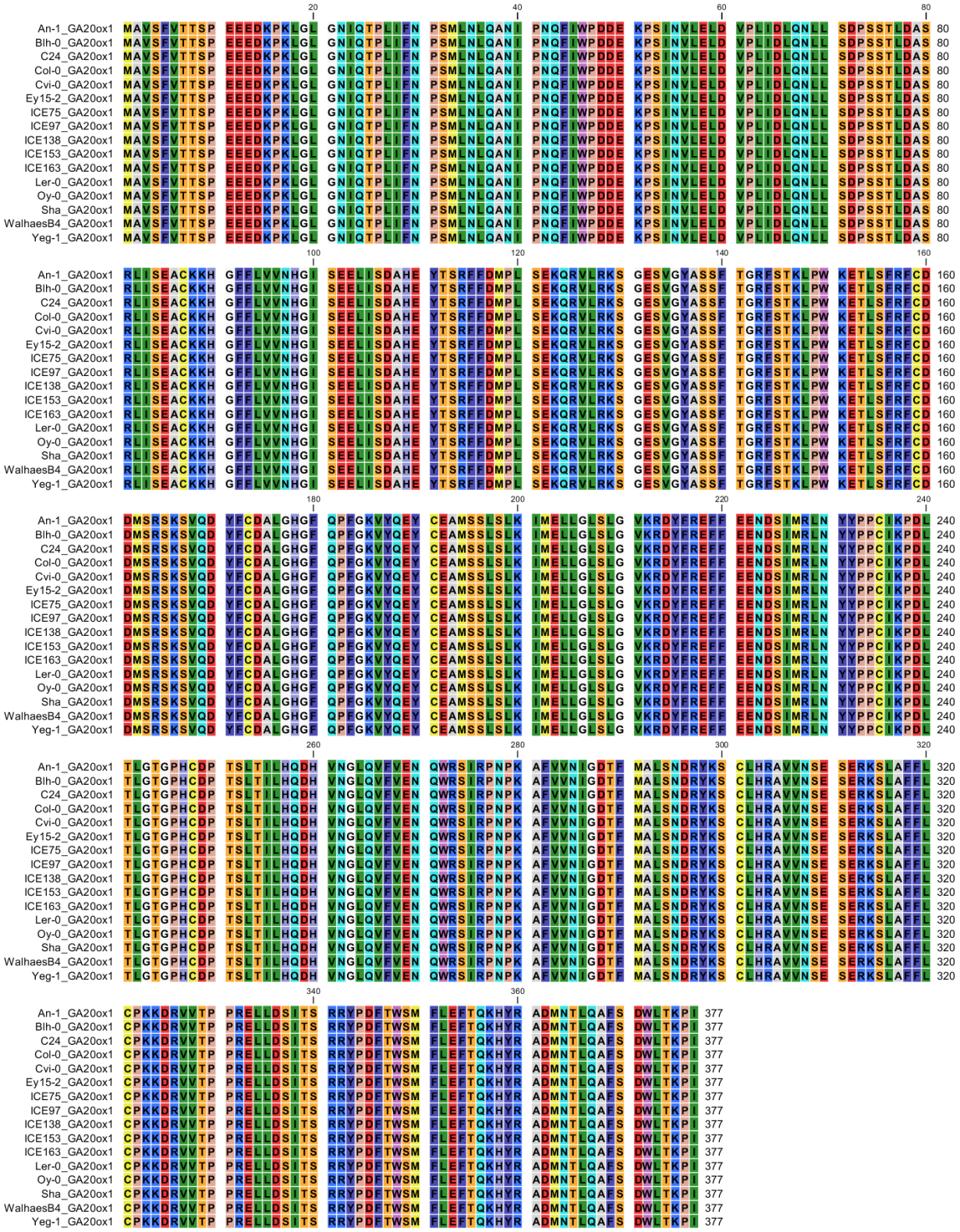
A


An-1_GA20ox1
Blh-1_GA20ox1
C24_GA20ox1
Col-0_GA20ox1
Cvi-0_GA20ox1
Ey15-2_GA20ox1
ICE75_GA20ox1
ICE97_GA20ox1
ICE138_GA20ox1
ICE153_GA20ox1
ICE163_GA20ox1
Ler-0_GA20ox1
Oy-0_GA20ox1
Sha_GA20ox1
WalhaesB4_GA20ox1
Yeg-1_GA20ox1

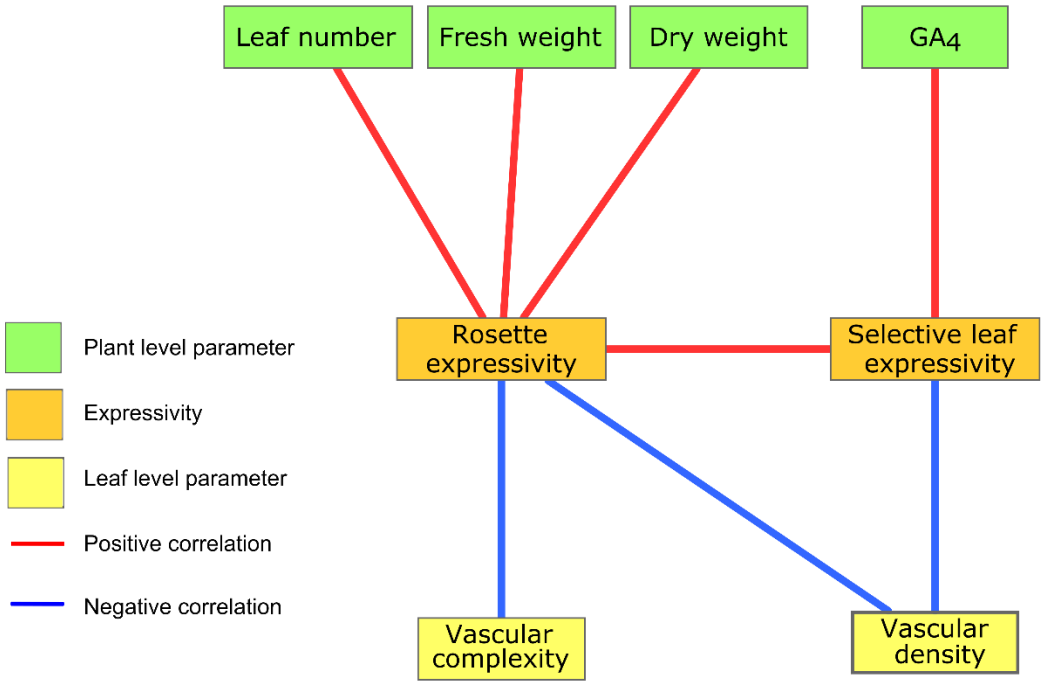

An-1_GA20ox1
Blh-1_GA20ox1
C24_GA20ox1
Col-0_GA20ox1
Cvi-0_GA20ox1
Ey15-2_GA20ox1
ICE75_GA20ox1
ICE97_GA20ox1
ICE138_GA20ox1
ICE153_GA20ox1
ICE163_GA20ox1
Ler-0_GA20ox1
Oy-0_GA20ox1
Sha_GA20ox1
WalhaesB4_GA20ox1
Yeg-1_GA20ox1


980
1.000
1.020
1.040
An-1_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
Blh-1_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
C24_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
Col-0_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
Cvi-0_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
Ey15-2_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
ICE75_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
ICE97_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
ICE138_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
ICE153_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
ICE163_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
Ler-0_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
Oy-0_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
Sha_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
WalhaesB4_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
Yeg-1_GA20ox1 TGTCCGAAAA AAGACAGAGT AGTGACGCCA CCGAGAGAGC TTTTGGACAG CATCACATCA AGAAGATACC CTGACTTCAC 1040
1.060
1.080
1.100
1.120
An-1_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
Blh-1_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
C24_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
Col-0_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
Cvi-0_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
Ey15-2_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
ICE75_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
ICE97_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
ICE138_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
ICE153_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
ICE163_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
Ler-0_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
Oy-0_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
Sha_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
WalhaesB4_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
Yeg-1_GA20ox1 ATGGTCTATG TTCCTTGAGT TCACTCAGAA ACATTATAGA GCAGACATGA ACACTCTCCA AGCCTTTTCA GATTGGCTCA 1120
An-1_GA20ox1 CCAAACCCAT CTAA 1134
Blh-1_GA20ox1 CCAAACCCAT CTAA 1134
C24_GA20ox1 CCAAACCCAT CTAA 1134
Col-0_GA20ox1 CCAAACCCAT CTAA 1134
Cvi-0_GA20ox1 CCAAACCCAT CTAA 1134
Ey15-2_GA20ox1 CCAAACCCAT CTAA 1134
ICE75_GA20ox1 CCAAACCCAT CTAA 1134
ICE97_GA20ox1 CCAAACCCAT CTAA 1134
ICE138_GA20ox1 CCAAACCCAT CTAA 1134
ICE153_GA20ox1 CCAAACCCAT CTAA 1134
ICE163_GA20ox1 CCAAACCCAT CTAA 1134
Ler-0_GA20ox1 CCAAACCCAT CTAA 1134
Oy-0_GA20ox1 CCAAACCCAT CTAA 1134
Sha_GA20ox1 CCAAACCCAT CTAA 1134
WalhaesB4_GA20ox1 CCAAACCCAT CTAA 1134
Yeg-1_GA20ox1 CCAAACCCAT CTAA 1134

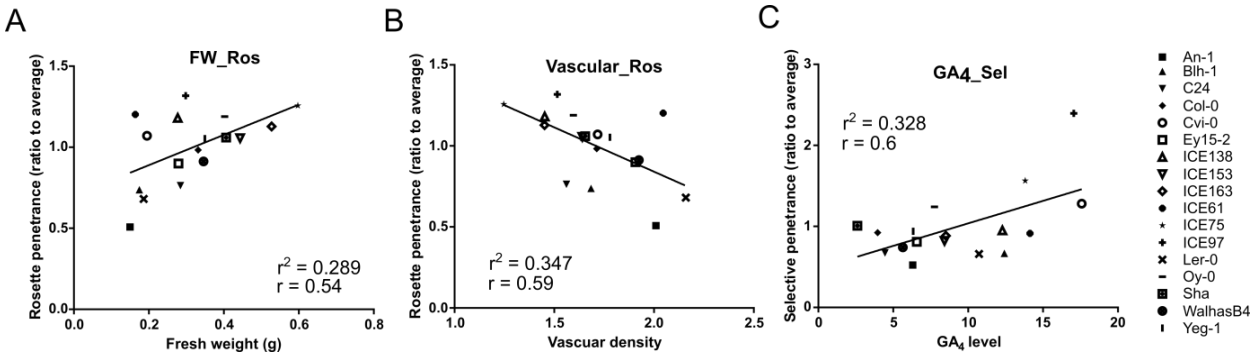
B



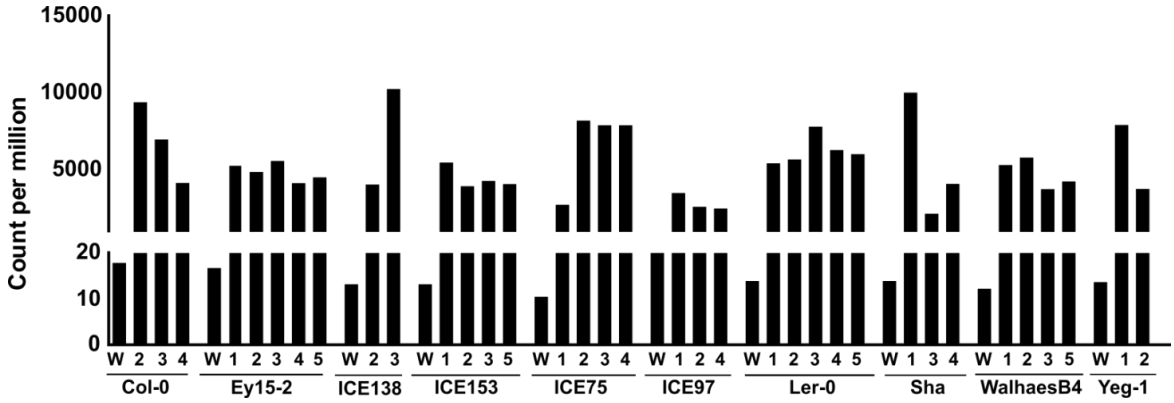
Extended Data Figure 5.5 Sequence alignments of the endogenous cDNA (A) and protein (B) sequence of GA20ox1 in 17 *Arabidopsis* accessions.



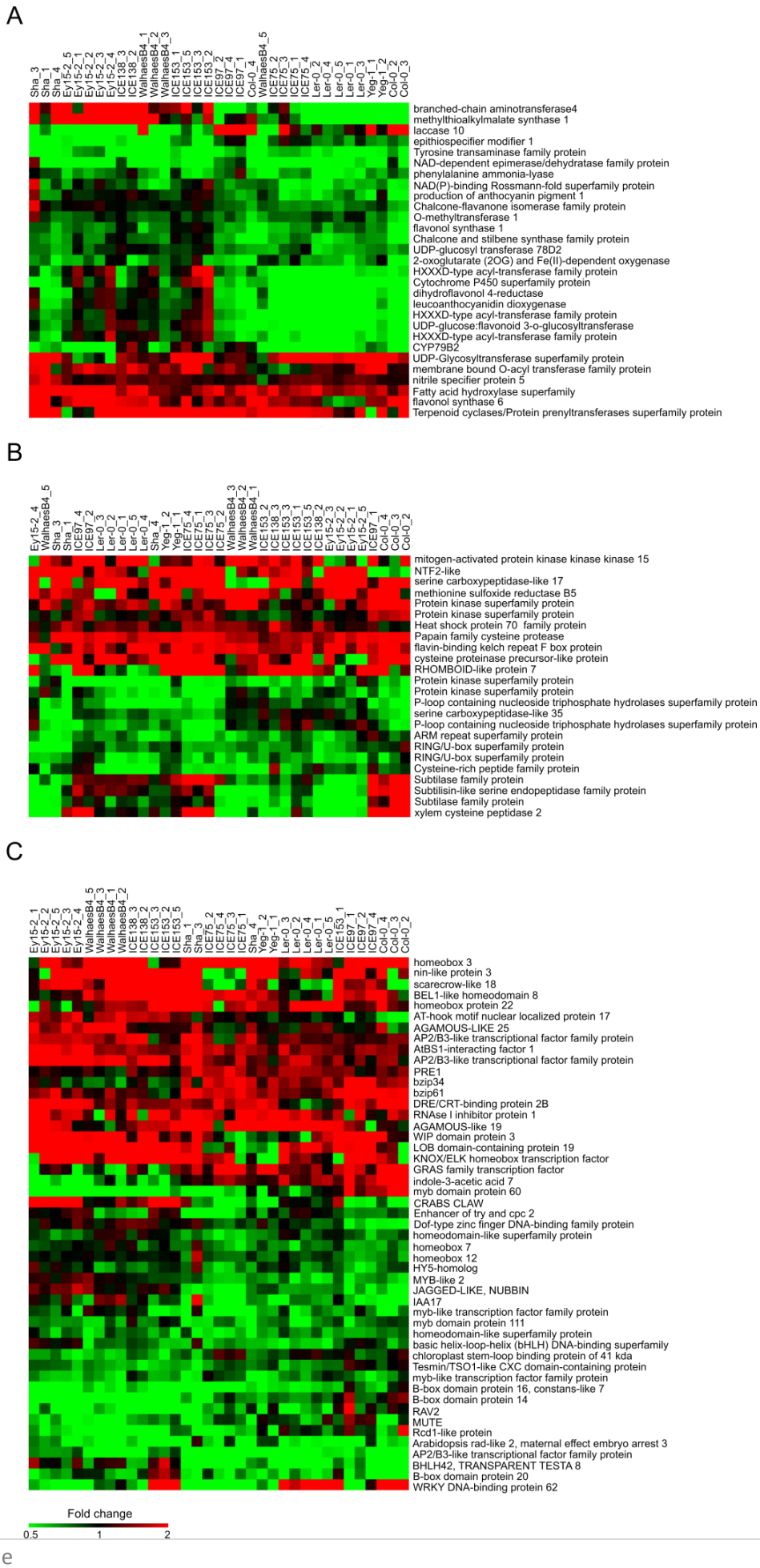
Extended Data Figure 5.6 Correlation analysis of phenotypic data. Correlation between phenotype parameters of wild types and expressivities of the *GA20ox1* ^{OE} effect in the transgenic lines. Node colours: green, plant level parameters; orange, expressivity; yellow, leaf level parameters. The red and blue edges show positive (correlation coefficient > 0.5) and negative correlation (correlation coefficient < -0.5) between parameters, respectively (adj-P value < 0.05).



Extended Data Figure 5.7 Scatter plots showing the correlation between the leaf size-related parameters in 17 Arabidopsis accessions and the expressivities. Scatter plots represent the correlation between (A) 'rosette' expressivity (Ros) and fresh weight (FW), (B) Ros and vascular density and (C) 'selective leaf' expressivity (Sel) and GA_4 levels in 17 accessions. A line representing linear regression is shown. R squared (r^2) and correlation coefficient (r) values are shown in the respective plots (adj-P value < 0.05).



Extended Data Figure 5.8 *GA20ox1* expression level in the transgenic lines from 10 Arabidopsis accessions. Absolute value (count per million) of expression level of *GA20ox1* in 10 wild types and transgenic lines from RNA sequencing data. W; wild type, 1-5; independent transgenic lines.


A
Sha_3
Sha_1
Sha_4
Ey15-2_5
Ey15-2_1
Ey15-2_2
Ey15-2_3
Ey15-2_4
ICE138_3
ICE138_2
WalhaesB4_1
WalhaesB4_2
WalhaesB4_3
ICE153_1
ICE153_5
ICE153_3
ICE153_2
ICE97_2
ICE97_4
ICE97_1
Col-0_4
WalhaesB4_5
ICE75_2
ICE75_3
ICE75_1
ICE75_4
Ler-0_2
Ler-0_4
Ler-0_5
Ler-0_1
Ler-0_3
Yeg-1_1
Yeg-1_2
Col-0_2
Col-0_3
branched-chain aminotransferase4
methylthioalkylmalate synthase 1
laccase 10
epithiospecifier modifier 1
Tyrosine transaminase family protein
NAD-dependent epimerase/dehydratase family protein
phenylalanine ammonia-lyase
NAD(P)-binding Rossmann-fold superfamily protein
production of anthocyanin pigment 1
Chalcone-flavanone isomerase family protein
O-methyltransferase 1
flavonol synthase 1
Chalcone and stilbene synthase family protein
UDP-glucosyl transferase 78D2
2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase
HXXXD-type acyl-transferase family protein
Cytochrome P450 superfamily protein
dihydroflavonol 4-reductase
leucoanthocyanidin dioxygenase
HXXXD-type acyl-transferase family protein
UDP-glucose:flavonoid 3-o-glucosyltransferase
HXXXD-type acyl-transferase family protein
CYP79B2
UDP-Glycosyltransferase superfamily protein
membrane bound O-acyl transferase family protein
nitrile specifier protein 5
Fatty acid hydroxylase superfamily
flavonol synthase 6
Terpenoid cyclases/Protein prenyltransferases superfamily protein
B
Ey15-2_4
WalhaesB4_5
Sha_3
Sha_1
ICE97_4
ICE97_2
Ler-0_3
Ler-0_2
Ler-0_1
Ler-0_5
Ler-0_4
Sha_4
Yeg-1_2
Yeg-1_1
ICE75_4
ICE75_1
ICE75_3
ICE75_2
WalhaesB4_3
WalhaesB4_2
WalhaesB4_1
ICE153_2
ICE138_3
ICE153_3
ICE153_1
ICE153_5
ICE138_2
Ey15-2_3
Ey15-2_2
Ey15-2_1
Ey15-2_5
ICE97_1
Col-0_4
Col-0_3
Col-0_2
mitogen-activated protein kinase kinase kinase 15
NTF2-like
serine carboxypeptidase-like 17
methionine sulfoxide reductase B5
Protein kinase superfamily protein
Protein kinase superfamily protein
Heat shock protein 70 family protein
Papain family cysteine protease
flavin-binding kelch repeat F box protein
cysteine proteinase precursor-like protein
RHOMBOID-like protein 7
Protein kinase superfamily protein
Protein kinase superfamily protein
P-loop containing nucleoside triphosphate hydrolases superfamily protein
serine carboxypeptidase-like 35
P-loop containing nucleoside triphosphate hydrolases superfamily protein
ARM repeat superfamily protein
RING/U-box superfamily protein
RING/U-box superfamily protein
Cysteine-rich peptide family protein
Subtilase family protein
Subtilisin-like serine endopeptidase family protein
Subtilase family protein
xylem cysteine peptidase 2
C
Ey15-2_1
Ey15-2_2
Ey15-2_5
Ey15-2_3
Ey15-2_4
WalhaesB4_5
WalhaesB4_3
WalhaesB4_1
WalhaesB4_2
ICE138_3
ICE138_2
ICE153_3
ICE153_2
ICE153_5
Sha_1
Sha_3
ICE75_2
ICE75_4
ICE75_3
ICE75_1
Sha_4
Yeg-1_1
Ler-0_3
Ler-0_2
Ler-0_4
Ler-0_1
Ler-0_5
ICE153_1
ICE97_1
ICE97_2
ICE97_4
Col-0_4
Col-0_3
Col-0_2
homeobox 3
nin-like protein 3
scarecrow-like 18
BEL1-like homeodomain 8
homeobox protein 22
AT-hook motif nuclear localized protein 17
AGAMOUS-LIKE 25
AP2/B3-like transcriptional factor family protein
AtBS1-interacting factor 1
AP2/B3-like transcriptional factor family protein
PRE1
bzip34
bzip61
DRE/CRT-binding protein 2B
RNAse I inhibitor protein 1
AGAMOUS-like 19
WIP domain protein 3
LOB domain-containing protein 19
KNOX/ELK homeobox transcription factor
GRAS family transcription factor
indole-3-acetic acid 7
myb domain protein 60
CRABS CLAW
Enhancer of try and cpc 2
Dof-type zinc finger DNA-binding family protein
homeodomain-like superfamily protein
homeobox 7
homeobox 12
HY5-homolog
MYB-like 2
JAGGED-LIKE, NUBBIN
IAA17
myb-like transcription factor family protein
myb domain protein 111
homeodomain-like superfamily protein
basic helix-loop-helix (bHLH) DNA-binding superfamily
chloroplast stem-loop binding protein of 41 kda
Tesmin/TSO1-like CXC domain-containing protein
myb-like transcription factor family protein
B-box domain protein 16, constans-like 7
B-box domain protein 14
RAV2
MUTE
Rcd1-like protein
Arabidopsis rad-like 2, maternal effect embryo arrest 3
AP2/B3-like transcriptional factor family protein
BHLH42, TRANSPARENT TESTA 8
B-box domain protein 20
WRKY DNA-binding protein 62
Fold change
0.5
1
2

Extended Data Figure 5.9 Heat maps representing the fold change of differentially expressed genes in *GA20ox1*^{OE} transgenics. Differentially expressed genes involved in secondary metabolism (A), protein synthesis (B), and regulation of transcription (C). Generic names of genes are shown on the right side of the heat map and sample names are indicated on the top of heat map. Red and green colours mean increased and decreased expression, respectively, in comparison with the wild types. Only DE genes that show at least a 1.5-fold change difference are shown. Hierarchical clustering was done for both genes and samples with Manhattan distance metrics.

Extended Table 5.1 Correlation between levels of different hormones in the 17 accessions. Pearson correlations between 36 biosynthesis, degradation intermediates and bioactive forms from 6 different hormones (GA, Auxin, Cytokinins, JA, SA and ABA) were calculated. Red color indicates positive correlation. (Correlation coefficient > 0.6, adj-P value < 0.05) tZ, *trans*-zeatin; tZR, tZ riboside; tZRPs, tZR phosphates; cZ, *cis*-zeatin; cZR, cZ riboside; cZRPs, cZR phosphates; DZ, dihydrozeatin; DZR, DZ riboside; DZRPs, DZR phosphates; iP, N^6-(Δ^2-isopentenyl)adenine; iPR, iP riboside; iPRPs, iPR phosphates; tZ7G, tZ-7-*N*-glucoside; tZ9G, tZ-9-*N*-glucoside; tZOG, tZ-*O*-glucoside; tZROG, tZ-R-*O*-glucoside; cZROG, cZ-R-*O*-glucoside; tZRPsOG, tZR phosphates-*O*-glucoside; cZRPsOG, cZR phosphate-*O*-glucoside; DZ9G, DZ-9-*N*-glucoside; iP7G, iP-7-*N*-glucoside; iP9G, iP-9-*N*-glucoside; IAAsp, indole-3-acetyl aspartic acid; IAIle + IALeu, indole-3-acetyl-*L*-isoleusine + indole-3-acetyl-*L*-leucine; IAPhe, indole-3-acetyl-*L*-phenylalanine.

		Cytokinins																				GA							ABA	SA	JA	Auxin			
		tZ	tZR	tZRPs	cZ	cZR	cZRPs	DZ	DZR	DZRPs	iP	iPR	iPRPs	tZ7G	tZ9G	tZOG	tZROG	cZROG	tZRPsOG	cZRPsOG	DZ9G	iP7G	iP9G	GA_4	GA_8	GA_9	GA_{19}	GA_{24}	GA_{44}	GA_{53}				IAA	IAAsp	IAIle + IALeu	IAPhe
Cytokinins	tZ																																			
	tZR	0.58																																		
	tZRPs	0.41	0.25																																	
	cZ	0.38	0.55	-0.31																																
	cZR	-0.03	0.46	-0.47	0.78																															
	cZRPs	-0.06	0.38	-0.44	0.84	0.88																														
	DZ	0.45	0.11	0.15	0.36	0.21	0.16																													
	DZR	-0.10	0.40	-0.32	0.67	0.80	0.81	0.40																												
	DZRPs	0.03	-0.14	0.62	-0.19	-0.24	-0.15	0.56	0.15																											
	iP	0.71	0.52	0.18	0.49	0.20	0.16	0.23	-0.04	-0.12																										
	iPR	0.19	0.66	-0.31	0.74	0.91	0.76	0.15	0.66	-0.33	0.47																									
	iPRPs	0.16	0.18	0.43	-0.18	-0.09	-0.26	-0.07	-0.28	0.08	0.58	0.21																								
	tZ7G	0.76	0.37	0.61	-0.02	-0.36	-0.35	0.32	-0.31	0.31	0.47	-0.16	0.19																							
	tZ9G	0.76	0.33	0.57	0.00	-0.33	-0.32	0.38	-0.28	0.34	0.48	-0.14	0.17	0.99																						
	tZOG	0.43	-0.23	0.06	-0.17	-0.37	-0.44	0.27	-0.39	-0.01	0.09	-0.28	-0.03	0.40	0.45																					
	tZROG	0.54	0.49	0.89	-0.07	-0.23	-0.22	0.18	-0.13	0.50	0.33	-0.04	0.36	0.71	0.71	0.07																				
	cZROG	-0.35	-0.03	-0.16	-0.16	0.07	0.01	-0.53	-0.01	-0.18	0.06	0.17	0.37	-0.32	-0.37	-0.35	-0.24																			
	tZRPsOG	0.52	0.31	0.83	-0.08	-0.28	-0.25	0.25	-0.17	0.50	0.20	-0.17	0.19	0.65	0.66	0.19	0.94	-0.43																		
	cZRPsOG	-0.28	0.17	-0.18	0.18	0.34	0.38	-0.43	0.30	-0.10	0.17	0.38	0.23	-0.26	-0.30	-0.55	-0.10	0.84	-0.27																	
	DZ9G	0.18	-0.16	0.07	0.12	-0.04	0.05	0.80	0.28	0.65	-0.05	-0.17	-0.31	0.41	0.47	0.26	0.12	-0.49	0.20	-0.35																
	iP7G	0.33	0.07	0.04	0.06	-0.07	-0.17	0.10	-0.28	-0.07	0.60	0.15	0.56	0.51	0.51	0.37	0.10	0.08	-0.01	0.07	0.14															
	iP9G	0.31	0.14	-0.03	0.16	0.00	-0.08	0.05	-0.23	-0.19	0.58	0.20	0.48	0.45	0.43	0.35	0.03	0.04	-0.06	0.04	0.08	0.97														
GA	GA_4	0.42	0.29	0.59	0.03	-0.30	-0.19	-0.21	-0.37	0.07	0.23	-0.24	0.06	0.53	0.46	0.03	0.56	-0.20	0.55	-0.03	-0.19	0.17	0.21													
	GA_8	0.51	0.52	0.17	0.20	-0.07	-0.04	-0.12	-0.23	-0.22	0.61	0.12	0.18	0.65	0.61	-0.12	0.36	-0.09	0.21	0.09	-0.01	0.49	0.50	0.53												
	GA_9	0.62	0.50	0.53	0.30	-0.11	0.01	0.05	-0.12	0.10	0.46	0.00	0.05	0.62	0.62	0.02	0.71	-0.42	0.73	-0.11	0.02	0.14	0.16	0.80	0.64											
	GA_{19}	0.53	0.48	0.27	0.36	0.02	0.07	0.12	-0.07	0.02	0.70	0.18	0.30	0.63	0.62	-0.07	0.44	-0.24	0.36	0.06	0.13	0.60	0.62	0.58	0.86	0.77										
	GA_{24}	0.32	0.21	-0.09	0.16	-0.07	-0.08	-0.02	-0.18	-0.24	0.70	0.15	0.42	0.34	0.33	0.03	0.00	0.12	-0.13	0.20	0.01	0.65	0.66	0.12	0.71	0.27	0.70									
	GA_{44}	0.19	0.26	0.49	0.15	-0.01	0.13	0.08	0.07	0.31	0.00	-0.09	-0.11	0.19	0.22	-0.20	0.57	-0.54	0.63	-0.28	0.02	-0.28	-0.28	0.53	0.16	0.69	0.35	-0.29								
	GA_{53}	0.36	0.28	-0.05	0.20	-0.03	-0.06	0.05	-0.14	-0.22	0.74	0.19	0.50	0.30	0.31	0.01	0.02	0.02	-0.12	0.06	-0.03	0.66	0.66	0.10	0.66	0.28	0.74	0.89	-0.07							
ABA		0.73	0.56	0.37	0.38	0.10	0.05	0.37	0.05	0.11	0.45	0.20	-0.02	0.74	0.75	0.24	0.64	-0.50	0.69	-0.19	0.30	0.30	0.32	0.55	0.57	0.80	0.69	0.25	0.38	0.19						
SA		-0.33	-0.12	-0.38	0.49	0.53	0.77	0.18	0.64	0.10	-0.20	0.32	-0.41	-0.44	-0.39	-0.34	-0.30	-0.12	-0.24	0.18	0.27	-0.35	-0.31	-0.28	-0.33	-0.15	-0.22	-0.22	0.06	-0.33	-0.19					
JA		0.58	0.22	0.50	0.23	-0.02	0.01	0.49	0.00	0.47	0.38	0.01	0.09	0.63	0.69	0.22	0.69	-0.45	0.78	-0.22	0.45	0.22	0.13	0.36	0.26	0.65	0.49	-0.05	0.58	0.05	0.73	-0.12				
Auxin	IAA	0.63	0.44	0.01	0.42	0.26	0.18	0.21	-0.04	-0.18	0.89	0.51	0.48	0.47	0.49	0.13	0.21	0.13	0.09	0.22	0.05	0.66	0.61	0.04	0.58	0.25	0.57	0.60	-0.17	0.63	0.42	-0.24	0.39			
	IAAsp	0.59	0.30	0.04	0.36	0.30	0.17	0.35	0.01	-0.03	0.69	0.47	0.38	0.43	0.46	0.33	0.21	0.10	0.14	0.13	0.15	0.60	0.52	-0.03	0.25	0.09	0.30	0.24	-0.19	0.28	0.40	-0.19	0.49	0.88		
	IAIle + IALeu	0.61	0.36	-0.08	0.46	0.37	0.22	0.38	0.08	-0.12	0.63	0.49	0.21	0.45	0.46	0.35	0.11	0.02	0.06	0.07	0.21	0.60	0.56	-0.01	0.30	0.10	0.34	0.23	-0.21	0.27	0.47	-0.20	0.46	0.84	0.96	
	IAPhe	0.53	0.28	0.04	0.36	0.32	0.17	0.37	0.03	0.02	0.59	0.44	0.31	0.45	0.48	0.31	0.21	0.05	0.16	0.10	0.22	0.62	0.54	0.01	0.25	0.10	0.32	0.16	-0.13	0.22	0.44	-0.21	0.55	0.82	0.97	0.97

Extended Data Table 5.2 Phenotype of *GA20ox1*^{OE} transgenics in 17 Arabidopsis accessions. Heat maps representing, per independent transgenic, the average percent difference in each leaf area between *GA20ox1* ^{OE} transgenics and their corresponding wild type (left) and the estimated expressivity (middle) of *GA20ox1* ^{OE}. More than 10 plants were analyzed. Fold changes of expression levels of *GA20ox1* compared to wild type analyzed by qPCR in T_2 generation are shown on the right panel (L1-L5; independent transgenic lines).

accession	Line	cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15	Sel	Ros	GA20ox1
An-1	L1	73	90	75	79	78											78	49	166
	L2	73	79	75	77	73											75	63	196
	L3	63	72	72	65	64											67	48	338
	L4	82	92	79	81	78											79	50	219
	L5	86	65	67	64	55											62	57	213
Blh-1	L1	94	99	100	104	92	90	101									98	76	221
	L2	90	90	82	93	94	100	112									87	82	234
	L3	97	112	103	107	99	102	110									104	84	309
	L4	89	103	92	92	90	84	84									95	65	213
	L5	84	89	84	93	87	97	109									86	82	205
C24	L1	82	87	92	97	107	98	104	102	109	83	70	65	52	44		73	91	1253
	L2	77	78	89	95	95	88	86	81	67	54	44	31	26	22		43	72	1517
	L3	66	77	90	92	94	97	94	91	80	61	58	36	28	23		52	77	1510
	L4	69	77	91	98	105	100	95	90	83	70	61	42	34	37		58	81	2237
	L5	82	94	93	100	104	102	98	89	83	69	52	38	32	35		53	82	1597
Col-0	L2	110	146	143	137	138	140	171									150	115	203
	L3	63	119	117	115	127	112	138									126	82	157
	L4	67	104	116	125	124	134	144									134	115	77
Cvi-0	L1	104	89	89	99	108	138	165	245	286	420						317	114	799
	L2	101	94	99	106	107	119	135	141	138	120						133	108	833
	L3	102	100	103	110	114	128	152	166	147	197						170	116	1098
Ey15-2	L1	99	96	100	102	106	103	109	108	131	124						121	89	500
	L2	101	101	97	94	96	105	112	109	122	118						116	101	510
	L3	100	99	95	93	99	99	105	104	115	116						112	92	704
	L4	97	100	98	98	103	106	104	108	120	115						114	94	704
	L5	96	100	99	96	107	101	115	114	137	133						128	98	430
ICE138	L2	97	107	105	108	126	134	129	122	122	117	113	105	103			130	118	140
	L3	89	103	108	113	133	148	159	138	162	157	145	162	165			147	131	199
ICE153	L1	91	97	98	98	103	111	101	111	110	106	112	87	81	81	97	109	106	50
	L2	90	96	95	94	101	111	103	104	103	95	99	87	75	69	57	99	100	101
	L3	90	95	99	96	104	113	108	115	112	112	113	92	91	75	76	112	107	126
	L5	102	109	121	112	128	132	127	139	156	158	184	155	194	170	166	166	132	174
ICE163	L1	83	107	102	100	102	114	119	138	131	146	148	125	147	143	136	139	124	282
	L2	105	116	110	111	121	135	148	159	158	168	158	156	151	156	146	161	143	74
	L3	74	87	88	87	104	107	108	105	90	91	82	66	67	58	59	95	92	29
	L4	72	86	85	93	102	112	120	123	124	116	112	99	93	88	81	121	109	78
	L5	82	97	103	106	117	138	141	145	138	137	141	120	119	115	112	140	127	58
ICE61	L1	110	131	128	124	121	136	139	145	154	143	149	146	125	116		146	135	646
	L2	109	121	132	119	127	133	127	126	123	117	111	100	98	70		126	123	541
	L3	105	116	125	120	124	129	129	125	132	117	105	100	87	71		129	122	1139
ICE75	L1	107	93	99	97	107	118	147	177	200	286	304	324	374	321	238	305	143	256
	L2	112	95	101	99	107	114	135	158	167	200	192	192	187	152	153	195	127	331
	L3	104	91	97	94	104	110	136	160	170	222	197	214	208	181	222	211	127	202
	L4	106	99	102	97	110	111	144	163	174	220	202	243	224	218	195	221	133	379
ICE97	L1	125	118	104	97	97	87	90	93	104	120	117	123	157	150		143	106	16
	L2	105	105	111	109	122	120	139	158	180	270	338	397	561	531		496	150	26
	L4	115	112	118	112	126	127	146	162	190	283	409	537	667	675		626	161	18
Ler-0	L1	82	80	77	77	90	96										80	64	318
	L2	87	85	90	94	111	134										87	67	236
	L3	62	66	67	64	78	87										65	72	37
	L4	81	92	86	91	107	126										86	81	329
	L5	78	81	86	89	99	113										82	74	197
OY-0	L1	101	98	95	101	108	117	128	135	138	152	180	177				170	120	77
	L2	88	97	96	100	117	123	137	138	151	154	166	175				165	132	131
	L3	94	91	90	95	104	114	126	133	153	155	204	208				189	126	96
	L4	107	96	97	101	105	116	127	139	153	174	206	232				204	124	121
Sha	L1	82	76	91	91	110	122	141	147								137	103	142
	L3	98	90	100	102	118	139	161	179								159	121	28
	L4	99	88	91	90	103	121	138	173								144	112	136
WalhaesB4	L1	91	105	110	119	115	110	108	105	101	84	73					115	107	253
	L2	73	81	90	90	102	83	66	53	33	29	21					92	72	264
	L3	80	102	97	105	112	116	118	112	107	98	71					111	99	272
	L5	90	102	104	106	117	119	114	118	113	104	96					114	108	299
Yeg-1	L1	112	123	114	112	113	103	110	100	116	130	145	158	162	171	100	164	118	33
	L2	88	109	96	102	98	90	93	97	110	123	107	108	121	116	91	115	104	61

Extended Data Table 5.3 Over-represented MapMan categories for *GA20ox1* DE genes. The number of genes found in each over-represented category is indicated. P-value with Bonferroni correction is shown.

Categories		Number of genes	P-value
Photosynthesis		46	5.14E-29
Secondary metabolism		47	3.98E-16
Protein		47	5.83E-11
Hormone metabolism		42	5.23E-10
RNA.regulation of transcription		92	
	RNA.regulation of transcription.C2C2(Zn) CO-like, Constans-like zinc finger family	8	1.33E-06
	RNA.regulation of transcription.MYB-related transcription factor family	8	3.60E-05
Transport		51	4.96E-05
Amino acid metabolism		19	8.55E-05
Sulfur-assimilation		3	
	Sulfur-assimilation.adenosine 5'-phosphosulfate reductase	3	2.02E-05

Chapter 6

Natural variation of morphological and molecular response to down-regulation of *PEAPOD2* in Arabidopsis

Youn-Jeong Nam[1,2], Dorota Herman[1,s2], Eunyoung Chae[3], Frederik Coppens[1,2], Veronique Storme[1,2], Twiggy Van Daele[1,2], Detlef Weigel[3], Dirk Inzé[1,2*] & Nathalie Gonzalez[1,2*]

[1]Department of Plant Systems Biology, VIB B-9052 Gent, Belgium, [2]Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Gent, Belgium, [3]Department of Molecular Biology, Max Planck Institute for Developmental Biology, Tübingen, Germany

* These authors contributed equally to this work.

Contributions: Y.J.N. was the main author of this work. Y.J.N., E.C, and T.V.D. conducted experimental work. D.H., F.C., V.S., and Y.J.N. analysed the data. D.I., N.G., and D.W. supervised the research. D.I. and N.G. contributed to the writing of this chapter.

Manuscript is in preparation and will be submitted for publication after minor changes to text.

Introduction

In Arabidopsis, after emergence of the leaf primodium from the shoot apical meristem, final leaf size is obtained by three interlinked events: a cell division phase, a cell expansion phase, and a meristemoid division phase. At least five different modifications of these events can positively influence final leaf size: an increased cell division rate, a prolonged period of proliferation, an increased cell expansion rate, a prolonged period of cell expansion, or a prolonged meristemoid division (Gonzalez et al., 2012).

Meristemoids are stem cell-like cells, dispersed in the leaf epidermis, that divide asymmetrically to give rise not only to a meristemoid cell but also to a pavement cell. In the epidermis, a meristemoid can undergo, at maximum, three sequential asymmetric divisions resulting in the generation of three pavement cells before differentiating into a stoma. Therefore, meristemoid division contributes to the production of a large portion of pavement cells (Geisler et al., 2000; Bergmann and Sack, 2007).

It has been shown that *PPD1/2* genes encoding transcription regulators belonging to the plant-specific TIFY protein family play an important role in the timing of meristemoid division. Arabidopsis has two classes of TIFY proteins based on the presence or absence of a C2C2-GATA domain, class l and class ll, respectively (Cuellar Perez et al., 2014). PPD proteins belong to class ll that contains also 12 JASMONATE ZIM DOMAIN (JAZ) proteins and TIFY8. TIFY proteins contain a highly conserved TIFY domain, also called ZIM domain, which mediates interactions between TIFY proteins (homo- and heteromeric dimerization) and with other specific transcription factors (Chini et al., 2009). Most of the class ll TIFY proteins have a C-terminal Jas domain known to mediate either binding to a bHLH MYC factor or the F-box protein COI1 (CORONATINE INSENSITIVE 1) depending on the absence or the presence of JA (Pauwels et al., 2010). The PPD proteins have, compared to the JAZ proteins, a divergent C-terminal Jas domain and an additional N-terminal PPD-domain (Bai et al., 2011).

The deletion or down-regulation of *PPD* genes induces enhanced leaf area and a dome-shape phenotype as a result of prolonged meristemoid division (White, 2006; Gonzalez et al., 2010; Gonzalez et al., 2015). This phenotype has been described in two genetic backgrounds, Ler-0 and Col-0 (White, 2006; Gonzalez et al., 2015). In Col-0, the effect of the down regulation of *PPD2* was studied at transcript level and the target genes and protein partners of PPD2 were identified (Gonzalez et al., 2015). Most of the differentially expressed genes in *amiPPD* were up-regulated consistent with the function of PPD as a negative transcription regulator. Genes encoding Cyclin D3 (CYCD3*)* were up-regulated in *amiPPD* transgenic lines and also found as direct target genes of PPD2. Since it has been speculated that CYCD3s positively regulate meristemoid division (Elsner et al., 2012; Lau et al., 2014), PPD2 might control the meristemoid activity via the transcriptional repression of *CYCD3* genes (Gonzalez et al., 2015).

Different genetic backgrounds can either restrict or pronounce the expression of a phenotype caused by a specific mutation or transgene, also referred as expressivity, and this genetic background effect has been shown in several model organisms (reviewed in Chandler et

al., 2013). Transferability of the effect of gene alteration is a substantial issue for crop yield improvement since obtaining stable phenotypes might depend on the genetic background. Therefore to better understand gene transferability of the effect of gene alteration, it is important to estimate the expressivity (the degree to which a genotype is phenotypically expressed in individuals) of a specific genetic perturbation in several naturally occurring genetic variants. Later, further study for identifying the molecular causes of the differences in expressivity can be done. To obtain a better insight into the effect of a transgene regulating leaf growth in natural variants, we analysed the growth phenotype as well as molecular response after silencing of *PPD* genes expression in 15 natural Arabidopsis accessions. Not all transgenics showed larger rosette leaves and also alteration of leaf size and shape varied between the accessions. We identified differentially expressed genes for *amiPPD* lines in the accessions showing common and accession-specific molecular responses depending on the genetic background. This work shows the importance of using natural variation as a tool to further characterise the role of PPD2 in the regulation of leaf growth.

Result

Phenotypic analysis of *amiPEAPOD (amiPPD)* transgenics in 15 accessions

We introduced an artificial micro RNA (amiRNA) targeting *PPD* (Parizotto et al., 2004) in 15 different natural Arabidopsis accessions including Col-0 and Ler-0 (Table 6.1). A twenty base pair modified sequence of *PPD* (targeting both *PPD1* and *PPD2*) from Col-0 was cloned to form a micro RNA which is under the control of the 35S CaMV promoter (Gonzalez et al., 2010) and introduced by floral dip (Clough and Bent, 1998) into 15 accessions. T_3 homozygote lines from 2 to 5 independent transformants of "the lowest *PPD* expressing" lines, still showing a range of expression, were selected having at least 30% of down regulation of PPD2 expression compared to wild type for each accession, (Extended Data Table 6.1) and grown in soil for 25 days. Representative pictures of the *amiPPD* transgenics in comparison to their respective wild types are shown in Figure 6.1A. Most transgenic lines have a dome-shape leaf phenotype (evaluated visually but also based on the number of cuts made on the leaf when doing leaf series) while transgenic lines of C24 and Sha do not show this morphological change (Figure 6.1A). To quantify leaf size in detail, individual leaf area of all transgenic lines was measured at 25 DAS by making leaf series. The average effect per accession and per independent transgenic line of the over-expression of *amiPPD* were then estimated. In general, most independent transgenic lines showed consistent effect within the same accession, however there were some exceptions for instance line 1 of Ler-0 and line 3 of ICE153 (Extended Data Table 6.1). We found, for most accessions, a significant increase in the area of the older leaves and a significant decrease for the younger leaves (Figure 6.1B and Extended Data Table 6.1). In two accessions, An-1 and Blh-1, *amiPPD* transgenic lines displayed rosette with almost all leaves larger than that of the corresponding wild type (Figure 6.1B and Extended Data Table 6.1). Interestingly, only transgenic lines of ICE163 showed almost no changes for most leaves compared to their corresponding control plants (Figure 6.1B and Extended Data Table 6.1). To compare how much the down-

regulation of *PPD* is affecting positively leaf growth in the 15 accessions, we estimated the expressivity of the transgene. Two different methods were applied to calculate this expressivity. First, because not all leaves are positively affected in the transgenic lines, leaf 3, 4 and 5 (the leaves showing the highest increased area compared to wild type in most accessions) were chosen for the calculation of the so called 'selective leaf' expressivity (Sel). This 'selective leaf' expressivity was determined by averaging the percentage to the wild type of the area of these three leaves (Figure 6.1C and Extended Data Table 6.1). Using this method of calculation, we found that down-regulation of *PPD* causes a clear positive effect on leaf area in most accessions as high as 141% in Blh-1 whereas in Cvi-0 and ICE163 no change was observed. Second, the expressivity at rosette level (Ros) was estimated by calculating the percentage to wild type of the total rosette area per accession (Figure 6.1C and Extended Data Table 6.1). At rosette level, a somewhat deviating effect of the down-regulation of *PPD* is observed in comparison to the 'selective leaf' expressivity. In 8 accessions including Cvi-0 and ICE163, the transgenics have a larger total rosette area. However, 5 accessions, Ey15-2, C24, Yeg-1, WalhaesB4 and ICE153, which show positive effect in the 'selective leaf' expressivity, gave either a negative outcome or no effect for the expressivity at rosette level.

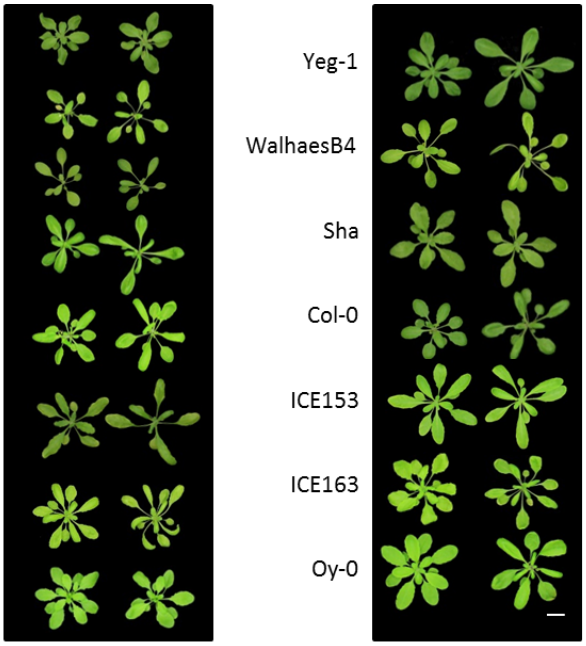
To examine if the accession-specific expressivity could result from differences in the sequence of *PPD2* between the accessions and Col-0, the cDNA and protein sequences of PPD2 were compared (Extended Data Figure 6.1). No sequence diversity at the target site of the artificial micro-RNA targeting *PPD2* was found between the accessions suggesting that the variation in expressivity does not seem to be related to the sequence variation at the target site of the artificial micro-RNA targeting of *PPD2*.

In conclusion, down-regulation of *PPD* causes alteration of either leaf shape and/or size depending on the accession suggesting that leaf shape and size are regulated in an independent manner. We found that only ICE163/*amiPPD* showed no change in leaf size but still an alteration of leaf shape.

Table 6.1 Geographic origin of the 15 Arabidopsis accessions used in this study.

Accession	Origin	Arabidopsis biological resource center stock number
An-1	Belgium	CS76091
Blh-1	Czech Republic	CS76098
C24	Portugal	CS76106
Col-0	Poland	CS76113
Cvi-0	Cape Verdi	CS76116
Ey15-2	Germany	CS76399
ICE138	Central Asia	CS76426
ICE153	Central Asia	CS76381
ICE163	Southern Tyrol	CS76353
ICE61	Russia	CS76378
Ler-0	Germany	CS77020
Oy-0	Norway	CS76203
Sha	Tadjikistan	CS76382
WalhaesB4	Germany	CS76408
Yeg-1	Kaukasus	CS76394

A



C

Accession	Sel	Ros
ICE61	128	110
An-1	137	152
Cvi-0	104	105
Ler-0	124	122
Blh-1	141	140
ICE138	125	110
Ey15-2	120	92
C24	112	83
Yeg-1	119	101
WalhaesB4	116	95
Sha	128	118
Col-0	135	116
ICE153	116	101
ICE163	99	97
Oy-0	126	105

B

Accession	cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15
ICE61	106	120	120	**125**	**137**	**138**	115	96	**81**	**74**	90	82	88		
An-1	**123**	**120**	**135**	**141**	**133**	**121**	100								
Cvi-0	107	110	101	100	99	84	**69**	**63**	**50**						
Ler-0	100	101	**121**	**124**	**124**	105	84	83							
Blh-1	**91**	**127**	**135**	**150**	**136**	**130**	**125**	**111**	94						
ICE138	108	118	**131**	128	114	106	88	**76**	**70**	**58**	**45**	**64**			
Ey15-2	105	117	120	118	108	112	109	94	83	**73**	**67**	**55**	**47**	**53**	**46**
C24	87	100	111	112	114	107	100	86	**63**	**48**	**36**	**28**	**26**	**22**	
Yeg-1	74	108	119	122	116	117	100	78	**64**	**44**	**42**	**31**	**36**		
WalhaesB4	96	**135**	**121**	**120**	106	102	95	**83**	**78**	**63**	**54**	**50**	**44**		
Sha	102	122	**132**	**129**	121	117	101	87	**66**	79	**64**				
Col-0	115	**129**	**134**	**137**	**132**	**122**	105	96	90	**77**	**69**				
ICE153	97	110	116	**118**	114	110	99	88	**78**	**65**	**65**	**55**			
ICE163	102	99	98	97	100	101	97	98	99	94	96	94	92	**85**	**81**
Oy-0	111	112	122	**128**	**128**	126	114	110	91	87	**78**	**66**	**57**	**61**	**40**
				60 <	60-80	80-90	90-95	95-105	105-110	110-120	120-140	> 140			

Figure 6.1 Phenotype of *amiPPD* transgenics in 15 Arabidopsis accessions. (A) Image of 25-days-old rosette (except Col-0, 24-days-old rosette) of representative *amiPPD* transgenics and their corresponding wild type. Scale bar: 2 cm. (B) Heat map representing, per accession, the average percent difference in leaf area between *amiPPD* transgenics and their corresponding wild type. Bold with underline: p-value < 0.05. (C) Heat map showing the estimated expressivity (Sel: 'selective leaf', Ros: 'rosette', see Methods) of *amiPPD*.

Cellular analysis in *amiPPD* transgenic lines

PPD is known to regulate the duration of meristemoid division (White, 2006) and a recent study showed that down-regulation of *PPD* causes increased cell number but no changes in cell area in a Col-0 background (Gonzalez et al., 2015). We tested whether different genetic backgrounds show distinct alterations or different levels of changes at cellular level upon down-regulation of *PPD*. Transgenic lines from An-1, Col-0, ICE163, Oy-0 and Sha were selected based on their leaf phenotype: high expressivity or low expressivity and dome-shape phenotypes or

not (Figure 6.1). Leaf 3 was chosen for the analysis since down-regulation of *PPD* has a positive effect in older leaves rather than younger leaves in most transgenic lines. Plants were grown for 25 days and the leaves were harvested to estimate pavement cell number and area (Figure 6.2).

Increased leaf 3 area was observed in the transgenic lines of An-1, Col-0, Oy-0 and Sha but not ICE163 in which a decreased area was found. All transgenic lines except ICE163 had an increase in pavement cell number of more than 20 % but the extent of the increase varied in function of the accession.

In conclusion, at cellular level, the effect of the down-regulation of *PPD* on leaf growth was mainly caused by changes in pavement cell number. In ICE163, the down-regulation of *PPD* seems however to not affect cell division.

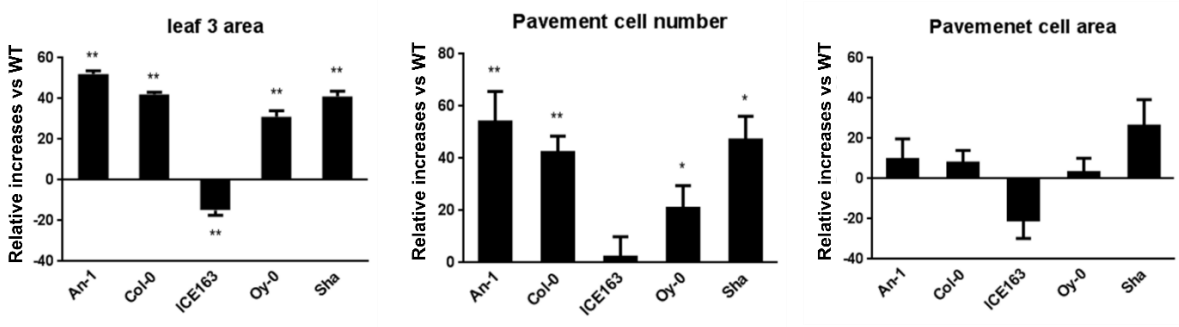


Figure 6.2 Cellular analysis of *amiPPD* transgenic lines in 5 different accessions. Five accessions (An-1, Col-0, ICE163, Oy-0 and Sha) were selected based on the expressivity of the transgenic lines. One transgenic line from each accession was chosen and grown in soil for 25 days. Leaves 3 were harvested and analyzed (3 leaves per transgenic and per repeat, in three biological repeats). Each graph shows the percentage of leaf 3 area, pavement cell number, and pavement cell area of the transgenic line compared to its corresponding wild type. Values represent average percentages to wild type with standard error. (ANOVA: * adj-P < 0.05, ** adj-P < 0.01)

Transcriptome analysis

In order to evaluate the molecular changes occurring upon down-regulation of *PPD,* transcript levels were analysed by RNA sequencing from 13 accessions (Ler-0, Col-0, An-1, WalhaesB4, Ey15-2, Yeg-1, Sha, Blh-1, ICE138, ICE163, Cvi-0, Oy-0 and C24). During leaf development, the transition phase between cell proliferation and cell expansion has been shown as a crucial stage to determine final leaf size (Andriankaja et al., 2012; Gonzalez et al., 2012). Also, meristemoid division occurs during cell division and cell expansion. Therefore at the transition phase, division of meristemoids is also present in the leaf. (Andriankaja et al., 2012). Thus, leaves were harvested for RNA sequencing analysis at the beginning of this transition phase, either at 9 or 10 DAS depending on the accessions (see methods).

Principal Component Analysis (PCA), conducted with the RNA-sequencing results from the 13 accessions showed a clear classification of the transcriptome data by accession (Figure

6.3). We confirmed the down-regulation of *PPD2* expression in all transgenic lines (Figure 6.4). However, *PPD1* expression, also targeted by the microRNA, showed almost no change in the transgenic lines. We found that the basal level of *PPD2* was higher than *PPD1* and variable between the accessions (Figure 6.4). Since no change in expression was found for *PPD1*, further analysis was mainly focusing on *PPD2* expression in the transgenic lines. Interestingly, we found a negative correlation between *PPD2* basal expression in wild type and its fold change in the transgenic lines showing that down regulation of *PPD2* is more pronounced in plants having high *PPD2* basal expression (Figure 6.4). For example, in Ler-0 which has low *PPD2* basal expression, the average fold change of *PPD2* expression in transgenic plants was 0.8 while transgenic plants of Yeg-1, the accession having high *PPD2* basal level, showed 0.6 fold change of *PPD2* expression.

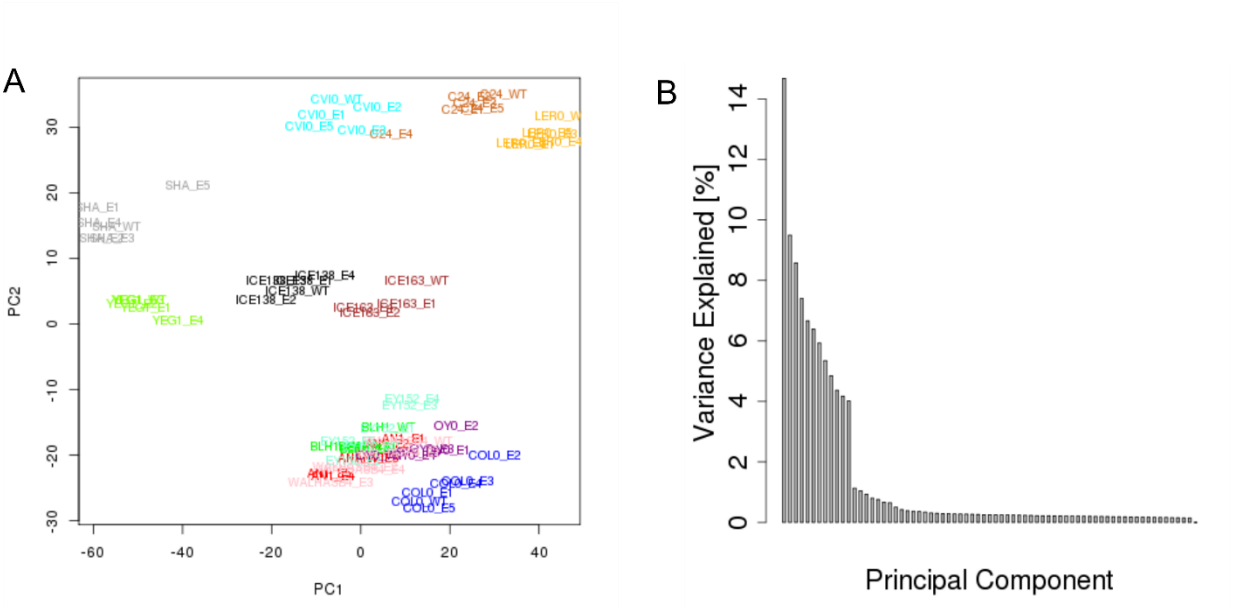


Figure 6.3 Principal Component Analysis (PCA) of the transcriptomic data. (A) The PCA plot shows the classification of the transcriptome data of wild type and *amiPPD* transgenics. Each accession is displayed in a different color. WT; wild type, E1-E5; independent transgenic lines. (B) Plot representing variance explained by PCA axes.

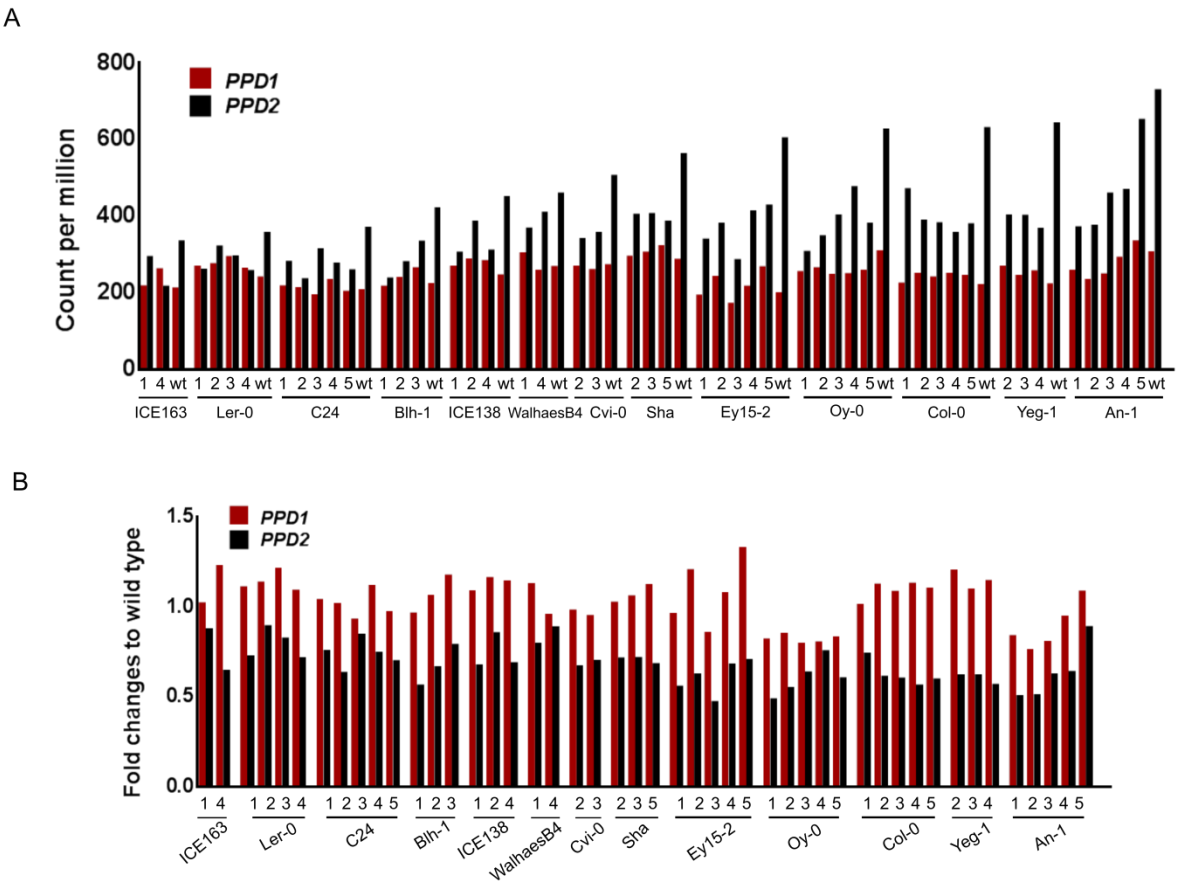


Figure 6.4 *PPD1* and *PPD2* expression level in the transgenic lines from 13 accessions. Absolute values from RNA sequencing data (count per million) of the expression levels of *PPD1* and *PPD2* in wild types and transgenic lines (A) and fold changes expression of *PPD1* and *PPD2* (B). The accessions are arranged based on the basal expression values of *PPD2* in the wild types. wt; wild type, 1-5; independent transgenic lines.

We then verified whether *PPD2* expression was negatively correlating with leaf area, but no correlation was found in the wild type with basal *PPD2* expression and in transgenics with down-regulation of *PPD2*. This result demonstrates that the level of down-regulation of *PPD2* might not be the only cause of the variation in the phenotypes of the transgenic lines, and that therefore other mechanisms are involved in making accessions responding differently to this genetic perturbation. Of course *PPD2* expression is measured in whole leaf samples while its molecular action is likely to be restricted to specific cell types and during specific developmental time points, rendering this type of correlation analysis difficult.

By considering changes in expression in at least one accession, 231 differentially expressed (DE) genes between wild type and *amiPPD* transgenic lines were identified. Overrepresented Gene Ontology (GO) categories were DNA, secondary metabolism, protein, hormone metabolism, mitochondrial electron transport/ATP synthesis, sulfur-assimilation, development, minor carbohydrate metabolism, signalling, and RNA related categories (Table

6.2). To gain more insight into the *amiPPD* DE genes, we performed a co-expression analysis with four predefined sub-sets of microarray expression data corresponding to four types of experiments (experiments in which leaf tissues are sampled (leaf), hormone treatment series (hormone), microarray experiments oriented toward growth, development and cell cycle studies (compendium1) or microarray experiments for which very similar experiments were removed (compendium2)) using CORNET (Correlation networks in plant) (De Bodt et al., 2010; De Bodt et al., 2012). Among the total number of 231 DE genes, we obtained networks with 61 genes which were present in the microarray expression data and showed high correlations with each other (correlation efficiency > 0.7). These genes are connected by 90 edges which can be divided in two main groups and some small networks (Figure 6.5). The first group contains 21 co-expressed genes connected with 57 edges that are involved in hormone metabolism, photosynthesis, cell wall and regulation of transcription. The second group contains 9 genes, linked with 13 edges, which are mainly associated with secondary metabolism. The other several small networks were formed by less than 5 genes each. We found two *CYCLIN* genes in one of the small networks, *CYCD3;2* and *CYCD3;3,* previously shown to be direct targets of PPD2 (Gonzalez et al., 2015). Therefore, down-regulation of *PPD2* leads to a change in expression of a set of genes that are co-regulated during leaf development.

Table 6.2 Over-represented Gene Ontology (GO) categories in the DE genes in the *amiPPD* lines. (P-value < 0.05).

GO categories	P-value
DNA	1.20E-04
secondary metabolism	2.13E-04
protein	7.96E-04
hormone metabolism	1.00E-03
mitochondrial electron transport / ATP synthesis	3.35E-03
S-assimilation	3.40E-03
development	5.40E-03
minor CHO metabolism	9.04E-03
signalling	0.039
RNA	0.033

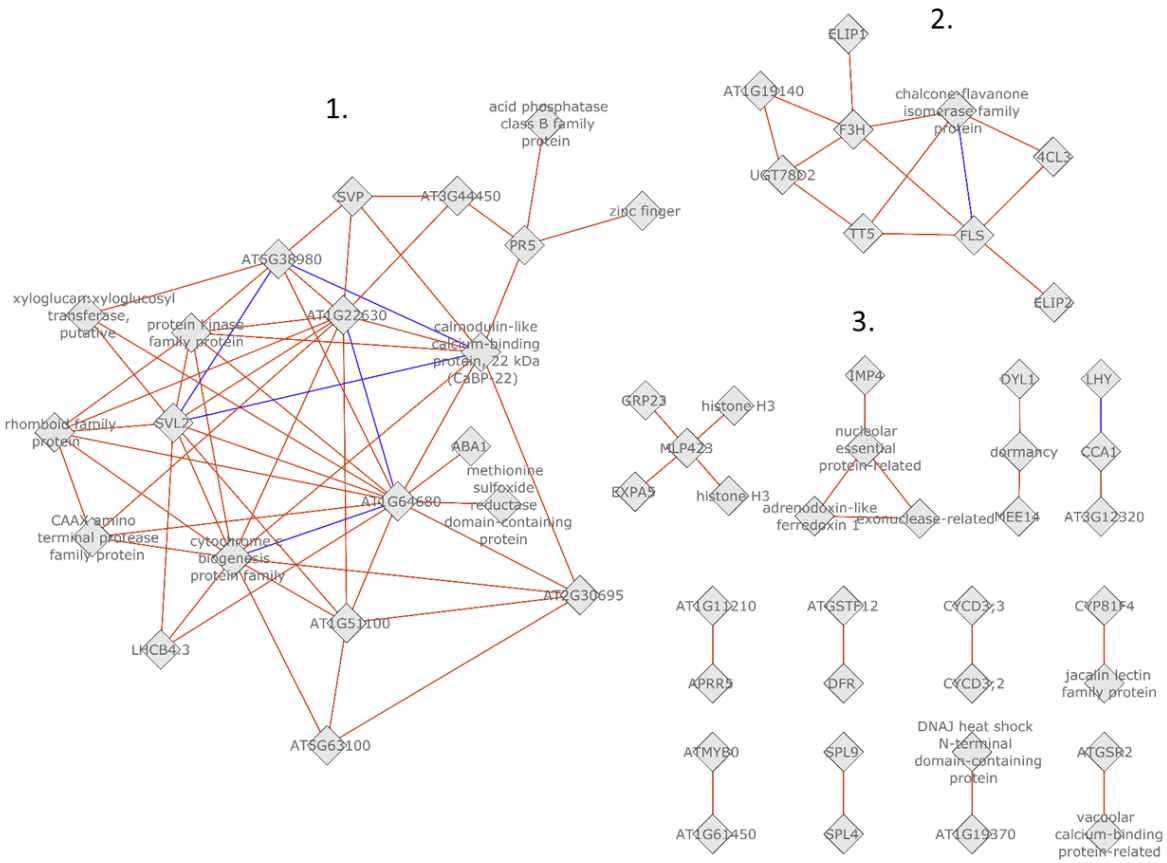


Figure 6.5 Co-expression networks with 61 genes from the DE genes in *amiPPD* transgenics. The network was made based on the Pearson correlation of the differential expression and divided in 3 different groups. (Edge color, orange: 0.7 < correlation coefficient < 0.8; dark blue: correlation coefficient > 0.8).

We compared the *amiPPD* DE genes in the 13 accessions to previously published microarray data from the first leaf pair of an *amiPPD* line in Col-0 harvested at 13DAS (time point at which the difference in leaf area starts to be visible) and a list of direct targets of PPD2 identified by tandem chromatin affinity purification-sequencing in Col-0 (Gonzalez et al., 2015)(Figure 6.6). Six genes out of the 49 genes DE in Col-0 were found in common and are up-regulated in most of the accessions (Figure 6.6A). For the comparison with the direct targets of PPD2, only genes bound by PPD2 in their 5' intergenic/untranslated regions (5' UTR) or promoter regions, corresponding to 918 genes, were used. We found 30 genes in common with 20 mostly up-regulated and 10 down-regulated in the majority of the transgenic lines (Figure 6.6B). Some genes related to DNA synthesis and amino acid metabolism showed down-regulation in the transgenic lines while genes encoding transcription factors and involved carbohydrate metabolism showed up-regulation. Some of these 30 genes also showed accession-specific pattern such as two genes encoding unknown proteins showing not only up-regulation but also down-regulation among the 13 accessions. Also two *Histone 3.1* genes (AT5G10390 and AT5G10400) showed strong down-regulation in the transgenic lines of ICE163 compared to other accessions (Figure 6.6B and Figure 6.7). We found that the basal expression level of the two

Histone 3.1 vary depending on the accessions (Figure 6.7). However, the basal levels of both genes in ICE163 were much lower compared to the other accessions.

In conclusion, we identified differentially expressed genes in *amiPPD* lines of 13 accessions and some of these genes were previously identified as PPD2 direct target genes in Col-0. We could observe somewhat accession-specific expression pattern for some target genes in the 13 accessions.

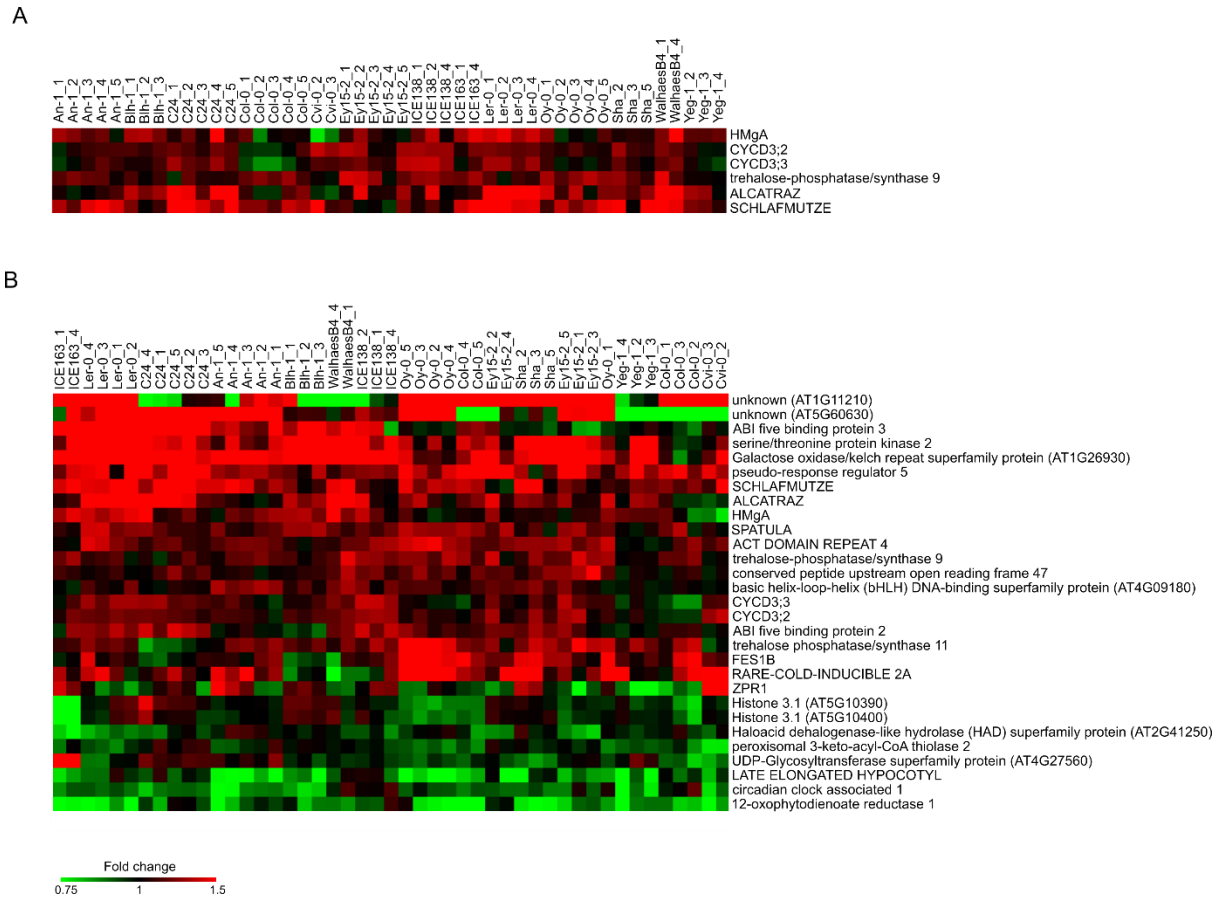


Figure 6.6 Heat maps representing the fold change expression of the differentially expressed genes in *amiPPD* transgenics overlapping with previously published data sets. (A) Differentially expressed genes overlapping with previously published microarray data obtained from the first leaf pair of *amiPPD* lines and wild-type plants in Col-0 harvested at 13 DAS (Gonzalez et al., 2015). (B) Differentially expressed genes overlapping with previously published tandem chromatin affinity purification sequencing (TChAP-seq) data corresponding to the direct targets of PPD2 (Gonzalez et al., 2015). Hierarchical clustering was done for both genes and samples with Manhattan distance metrics. Red color represents increased expression and green color decreased expression in comparison to the wild types.

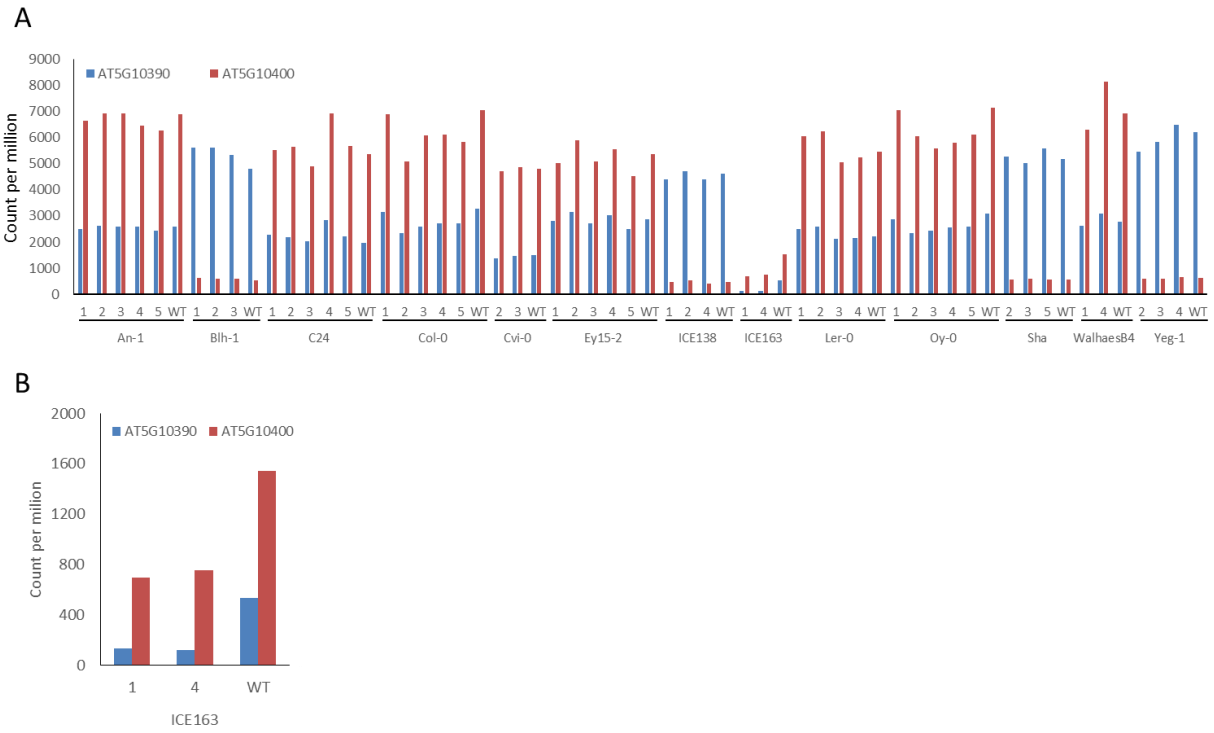


Figure 6.7 *HISTONE 3.1* expression levels in the *amiPPD* transgenic lines from the 13 accessions. Absolute values from the RNA sequencing data (count per million) for two *HISTONE 3.1* in 13 wild types and the transgenic lines (A). Expression level of two *HISTONE 3.1* genes (AT5G10390 and AT5G10400) that were differentially expressed in *amiPPD* transgenic lines of ICE163 (B). WT; wild type, 1-5; independent transgenic lines.

Correlation analysis between phenotypic and transcriptome data

Next we examined whether the effect of the down-regulation of *PPD*, previously defined as 'selective leaf' expressivity and 'rosette' expressivity (Figure 6.1) could be related to the phenotype of the wild-type accessions (data from chapter 5) and/or the differentially expressed genes. First, we found that the two expressivity values were highly correlated (Figure 6.8). We also found that both rosette and 'selective leaf' expressivity values were negatively correlating with fresh weight, dry weight and leaf number of the wild type accessions (Figure 6.8). Additionally rosette area of wild type accessions was negatively correlated with 'rosette' expressivity.

To assess whether molecular changes are related with high or low expressivity correlation analysis between differentially expressed genes and the two expressivity values was done. Only three differentially expressed genes correlating with 'rosette' expressivity were identified (Figure 6.9A). A gene encoding KUNITZ TRYPSIN INHIBITOR1 showed a clear negative correlation with 'rosette' expressivity while two other genes encoding DORMANCY ASSOCIATED GENE2 and TARGETING PROTEIN FOR XKLP2 PROTEIN FAMILY showed positive correlations. We also identified five DE genes correlating with *PPD2* down-regulation in the transgenic lines

(Figure 6.9B). Interestingly, these genes are positively correlated with *PPD2* expression. When *PPD2* is less down-regulated, these genes show up-regulation.

In conclusion, an interesting correlation was found between the phenotypes of wild type and expressivity after down-regulation of *PPD*. We also identified differentially expressed genes that are correlated with expressivity and *PPD2* down-regulation.

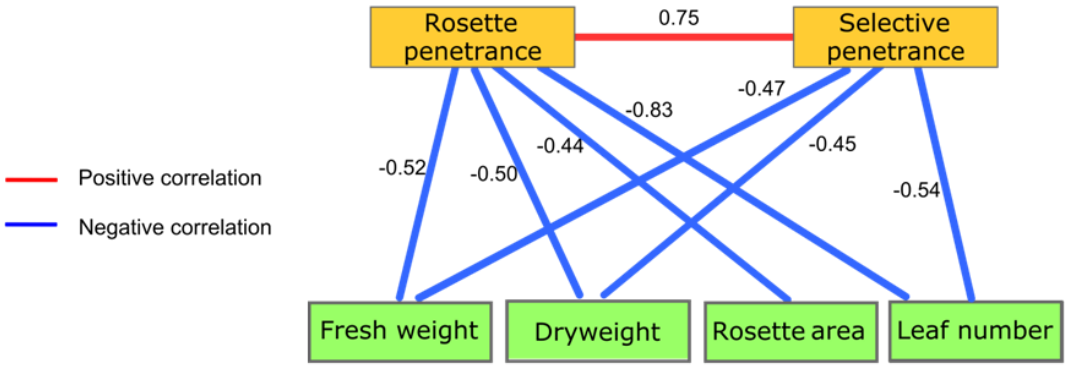


Figure 6.8 Correlation analysis between the phenotypic data. Correlation between different phenotypic parameters of wild types and the expressivity values of *amiPPD*. Node color, green; growth-related parameters, Orange; Expressivity. The red and blue edges show positive and negative correlations between the parameters (adj-P value < 0.05). Correlation coefficient values corresponding to the edges are shown.

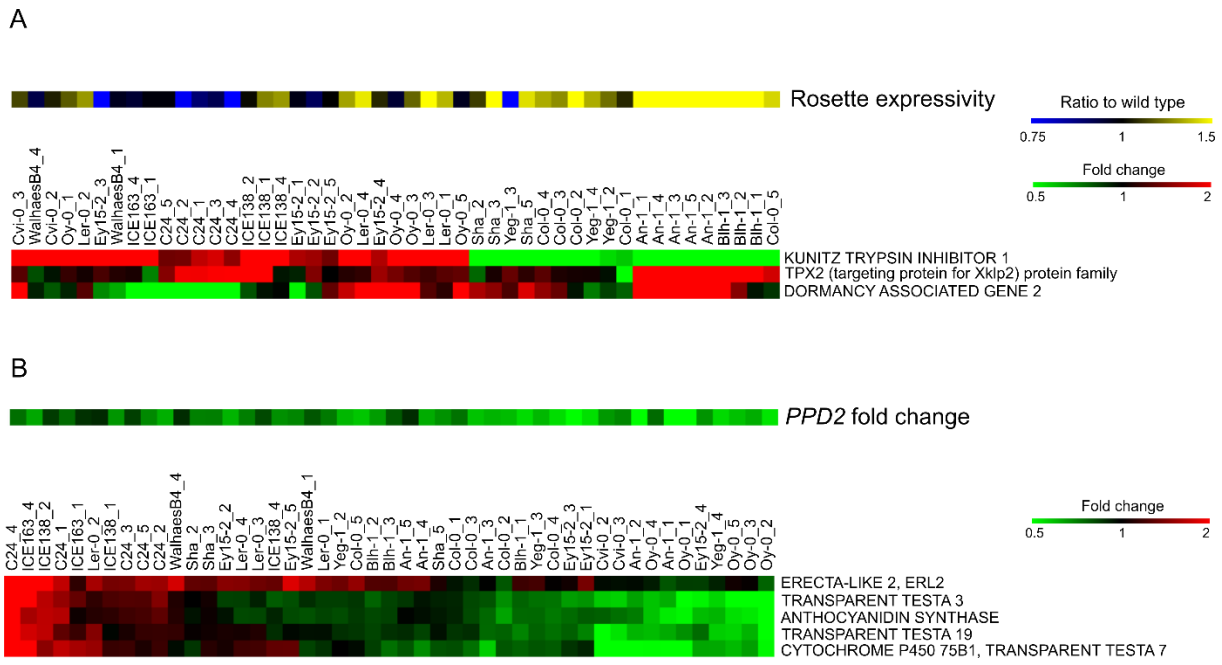


Figure 6.9 Correlation analysis between phenotypic and transcriptome data. Heat maps representing the DE genes correlated with 'rosette' expressivity values (A) and fold change of *PPD2* (B). For each panel, the upper heat maps show the 'rosette' expressivity and fold change of *PPD2*, respectively. Lower heat maps represent DE genes correlated with 'rosette' expressivity and fold change of *PPD2*, respectively. Red and green colors correspond to increased and decreased expression in comparison to the wild types, respectively. Hierarchical clustering was done for both genes and samples with Manhattan distance metrics. (Correlation coefficient > |0.5|, adj-P value < 0.05).

Discussion

In order to analyze the effect of the down-regulation of *PPD* in different genetic backgrounds, we introduced an *amiPPD* overexpressing construct in 15 Arabidopsis accessions. Previous studies have shown that in both the Ler-0 and Col-0 background, down-regulation of *PPD* causes the formation of larger leaves with a dome-shape phenotype (White, 2006; Gonzalez et al., 2015) and this increase in size results from an increased cell number (Gonzalez et al., 2015). In our study, the majority of the transgenic lines had larger older leaves except Cvi-0 and ICE163 but not all transgenic lines showed a larger rosette area (7 out of 15). Moreover, transgenic lines in C24 and Sha showed an increased leaf area without having any an obvious dome-shape phenotype. We also observed, in ICE163 background, that transgenic lines displayed a dome-shape leaf phenotype without having an increased leaf area. These results suggest that changes in leaf shape and area might be regulated in an independent manner. In ICE163 background, no change in pavement cell number but a decrease in pavement cell area, leading to a decreased leaf area, was observed. If PPD2 influences leaf size and shape separately by regulating different targets, it might be that in ICE163 down-regulation of *PPD2* only affects the targets for leaf shape not for size. Furthermore, we observed different degrees for increased pavement cell number in different accessions. Probably, down-regulation of PPD2 affects differentially, in the various accessions, the time window during which meristemoid division occurs, resulting in different effects on the number of pavement cells.

Through correlation analysis, we discovered that the effect of the down-regulation of *PPD2* is negatively correlated with biomass, rosette leaf area and leaf number of wild types. Indeed, we observed high 'rosette' expressivity, thus an increased rosette area compared to wild type, in An-1, Ler-0 and Blh-1 accessions that have less leaves whereas the Ey15-2 and C24 accessions producing more leaves showed low 'rosette' expressivity. Since wild-type accessions produce different number of leaves, it probably means that they have different timing for vegetative development as observed by the difference in the onset of flowering in the accessions. Thus, younger leaves from the accessions producing less leaves might be close to maturity while younger leaves from accessions producing more leaves might still be growing so did not reach their final size. In the latter case the effect of down-regulation of *PPD2* might not have been visible because the harvesting time was too early for these leaves. This could result in having no increased rosette area in these accessions. However, because of some early flowering accessions (An-1, Ler-0, and Blh-1) which start to flower before 25 DAS, thus, 25DAS was best time point to harvest leaves.

We identified 231 DE genes upon down-regulation of *PPD2* and only 6 of these genes were found in common with previously published microarray data set from Col-0/*amiPPD*. The use of different time points for the transcriptome analysis might be a reason to only find few genes in common between these data. Among the 231 DE genes, 30 of them were previously found to be direct targets of PPD2 (Gonzalez et al, 2015). Two third of these target genes were up-regulated while the others down-regulated in most transgenic lines across the different

accessions. From the up-regulated gene set, two cell cycle related genes *CYCD3;2* and *CYCD3;3* were found. Interestingly, *CYCD3;2* is known to be a direct target of SPEECHLESS (SPCH), a transcription factor that positively regulates meristemoid division, and is up-regulated upon induction of *SPCH* (Lau et al., 2014). Furthermore, it has been shown that the expression of both *CYCD3;2* and *CYCD3;3* is repressed by PPD2 and suggested that by repressing the expression of those two *CYCD3* genes PPD2 influences meristemoid activity (Gonzalez et al., 2015). However, the degree of changes in expression these genes varies between accessions. Another gene differentially expressed, also direct target of PPD2, is *HMGA* encoding a member of the chromatin-associated high mobility group (HMG) protein. HMGA is known to be regulated by cyclin-dependent kinase via phosphorylation (Zhao et al., 2009). In animals, HMGAs function as highly connected central nodes of protein-DNA and protein-protein interactions therefore acting as architectural transcription factors that regulate diverse biological processes including growth, proliferation, differentiation and death (Reeves, 2001). Loss of function mutant of *HMGA* had a dwarf phenotype due to decreased rate of cell proliferation in mouse (Zhou et al., 1995) and chicken (Ruyter-Spira et al., 1998). The expression levels of plant *HMGA* genes possibly correlated to the proliferative state of the cells (Gupta et al., 1997; Zhang et al., 2003). In maize, HMGA is up-regulated during endoreduplication in vegetative tissue suggesting that HMGA plays a role in the formation of chromatin relaxation to enable transcription and/or replication (Zhao and Grafi, 2000). It has been suggested that PPD might inhibit cell proliferation by repressing expression of HMGA which could reshape the chromatin organization (Gonzalez et al., 2015). Interestingly, in the list of 30 genes previously found to be direct targets of PPD2 we found two *HISTONE 3.1* genes encoding chromosome architectural proteins, which are known to interact with HMGA in animal (Reeves and Nissen, 1993). In leafy spurge, *HISTONE 3* is highly expressed in shoot apical meristem and it is up-regulated after treatments with gibberellin known to induce S-phase of cell cycle (Horvath et al., 2002; Horvath et al., 2003). Also cold and drought stress caused down-regulation of *HISTONE 3* suggesting that *HISTONE 3* is responsive to growth-regulating signals (Horvath et al., 2003). In Arabidopsis, *HISTONE 3.1* might also function as positive regulator of cell division. We observed a strong down-regulation of these genes in transgenic lines of ICE163 which did not show increased leaf area. Since other accessions such as Oy-0 and Col-0 having increased leaf area showed down-regulation of these genes, thus, it might be that a threshold expression level of these genes pushes the cell divisions. Also it could be that these genes are regulated in post-transcriptional manner. Furthermore, basal expression of these genes was extremely low in ICE163 wild type showing relatively high biomass of plant compared to other accessions. There might be a different mechanism to regulate these genes and to control the leaf size in ICE163. For example, other positive regulators for leaf growth are more active in this accessions.

We found that some of the direct targets of PPD2 were down-regulated in *amiPPD* transgenic lines. PPD is one member of the plant-specific TIFY transcription factor family (Bai et al., 2011; Cuellar Perez et al., 2014) which interacts with some of the TIFY proteins through a tify domain (Gonzalez et al., 2015). It has been shown that tify domain could act as transcriptional

activator (Shikata et al., 2003). In addition, two proteins that form a protein complex with PPD2 are KIX8 and KIX9 and contain a KIX protein domain (Gonzalez et al., 2015). The KIX protein domain has been shown to be present in several transcriptional co-activators (Radhakrishnan et al., 1997). Therefore, it can be speculated that PPD2 could act also as transcription activator through interaction with co-activators. To verify the hypothesis that PPD2 activates the expression of some target genes, transient expression assay in protoplasts with promoter-luciferase reporter constructs can be conducted. In addition, *PPD2* overexpression lines could be generated in ICE163 and Oy-0 accessions, both showing down-regulation of many target genes of PPD2 compared to other accessions to check if the down-regulated target genes in *amiPPD* lines would be up-regulated.

The correlation analysis identified DE genes that are correlated with the 'rosette' expressivity and *PPD2* down-regulation. Some of these genes are already known to be related to cell division or leaf growth. For example, *TPX2 (TARGETING PROTEIN FOR XKLP2)* and *DORMANCY/AUXIN ASSOCIATED GENE 2* are positively correlated with 'rosette' expressivity. Interestingly, *TPX2* is known to be positively involved in cell division control by regulating mitotic spindle organization (Vos et al., 2008). Also expression of AtDRM2 is responsive to a wide range of abiotic, physical and hormonal treatments (Rae et al., 2014). The expression of *ERECTA-LIKE 2,* a member of the *ERECTA* family*,* was found to have a positive correlation with *PPD2* expression in the transgenic lines. This gene showed up-regulation in the transgenic lines having less reduced *PPD2* expression, however, there was no correlation with the level of down regulation. Especially, transgenic lines in C24, ICE163, and ICE138 showed up-regulation of this gene. Since members of the *ERECTA* family are known to promote cell proliferation during organ growth (Shpak et al., 2004), probably some accessions have a different feedback mechanism that could buffer down-regulation of *PPD2*.

We observed that direct targets of PPD2 showed accession-specific differential expression between the accessions. It is possible that the level of PPD2 protein is differentially regulated in different accessions which leads to different activities of PPD2. This differential gene regulation could be also explained by epistatic interactions with other transcription regulators having the same target genes as PPD2 and which could have accession-specific expression. Depending on the activity of these transcription regulators in different accessions, the expression of the target genes will vary. We also observed that differential expression is not always in the same direction, for example, in some accessions, some of PPD2 target genes were up-regulated while in other accessions they showed down-regulation. This also could be caused by more complex epistatic events between the genetic background and down-regulation of *PPD2*.

We found that down-regulation of PPD2 affects leaf shape and size while in specific cases only leaf size or shape is affected. To identify molecular mechanisms that affect either leaf shape or size in different transgenic lines, additional comparative transcriptome analysis can be conducted in specific accessions. For example it would be of interest to analyse in detail the

effect of PPD down-regulation in Sha (only an increased leaf area), ICE163 (only modified leaf shape but no effect in leaf size) and Col-0 in which both size and shape effects are observed.

Here we demonstrated that the effect of down-regulation of *PPD2* differs depending on the genetic background. The alteration of leaf shape and leaf area caused by down-regulation of *PPD2* seems to be regulated by independent mechanisms. We also identified DE genes in the different accessions that are also direct targets of PPD2. Further detailed analysis on specific accessions could help to identify specific regulatory gene network regulated by PPD2.

Methods

Plant material and growth conditions

Fifteen Arabidopsis accessions were selected to cover most common genetic variation of *Arabidopsis thaliana* (Table 1) and used to generate lines overexpressing *amiPPD*. A twenty base pair sequence of *PPD* (targeting both *PPD1* and *PPD2*) from Col-0 was cloned under the control of the 35S CAMV promoter in the fluorescence-accumulating seed technology (FAST) vectors (Shimada et al., 2010) and introduced into the 15 accessions following the floral dip protocol (Clough and Bent, 1998). Dried transgenic T_1 seeds were selected based on fluorescence signal in the seed coat and sown on soil for seed production. T_2 transgenic seeds were harvested and selection of five independent single-locus insertion lines (75% of fluorescent seeds) was done. Seeds were sown on soil for seed production and expression of the transgene was verified by RT-qPCR. From these lines, at least two and maximum five independent T_3 homozygote lines for each accession were selected for further experiments. All plants were grown MIRGIS platform in soil under a 16-hours-day/8-hours-night regime at 21°C in a growth chamber. The daily images of plants obtained from MIRGIS are analyzed in Chapter 8.

Phenotypic analysis

Leaf series data were obtained from 25-day-old plants by dissecting individual leaves (from cotyledons to the younger rosette leaves) and mounting them on a 1 % agar plate, and the leaf area was measured with the ImageJ software (http://rsb.info.nih.gov/ij/). The leaf series for the transgenic lines were performed in one biological repeat by growing the transgenic lines together with their corresponding wild type.

For the leaf series data, leaf area was log transformed to stabilize the variance. The mean model consisted of the main effects of *amiPPD* on leaf size and their interaction term. Due to the unbalanced and complex nature of the data, the Kenward-Rogers approximation for computing the denominator degrees of freedom for the tests of fixed effects was used. An autoregressive structure was used to model the correlations between measurements done on the leaves originating from the same plant. The main interest was in the effect of the gene on leaf area for each leaf separately. Simple tests of effects were performed at each leaf between the transgenic lines and the corresponding wild type. Difference estimates were represented as %

to the least-square means estimate of the wild type and leaf. Separate models were run for each accession as they were grown in separate experiments. For each experiment, the data was truncated so that there were at least two observations for each leaf of both the transgenic lines and the corresponding wild type. The analysis was performed with the mixed and plm procedure of SAS [Version 9.4 of the SAS System for windows 7 64bit. Copyright © 2002-2012 SAS Institute Inc. Cary, NC, USA (www.sas.com). Residual diagnostics were carefully examined.

The expressivity of 'selective leaf' was determined by averaging the ratio of leaf 3, 4 and 5 for each leaf. To calculate the 'selective leaf' expressivity per accession, the 'selective leaf' expressivity per transgenic line were taken for the average estimated expressivity. 'Rosette' expressivity was estimated as a ratio of the wild-type rosette area to that of a transgenic line. In case of 'rosette' expressivity per accession, the mean of 'rosette' expressivity per transgenic line of an accession has been taken.

Cellular analysis was done as previously described (Andriankaja et al., 2012) in three biological repeats. For transgenic lines, significant differences were determined with a two-way ANOVA test with genotype and repeat as main factors in R software (v 3.0.1) (R Core Team, 2015). Differences between the wild type and corresponding transgenic lines were estimated and declared significant when the adj-P value < 0.05 with Tukey's method.

RNA extraction

Total RNA was extracted from the shoot of 12-day-old seedlings of T_2 transgenic lines and the corresponding wild-type plants using TRIzol and the expression of the transgene was analysed by Quantitative reverse transcription-PCR (RT-qPCR). RT-qPCR was performed as previously described (Claeys et al., 2012).

For RNA sequencing (RNA-seq) analysis, seedlings with one biological repeat of wild-type plants and overexpressing-*amiPPD* lines (at 8 DAS Yeg-1 and Sha and at 9 DAS for WalhaesB4, Col-0, ICE138, Blh-1, Ey15-2 and ICE163 and at 10 DAS Ler-0, An-1, Oy-0, Cvi-0 and C24) were harvested in RNA ice-later solution (AM7030; Ambion) and incubated at -20°C for at least a week. Leaf 3 was micro-dissected on a cold plate with dry ice under a stereomicroscope, and frozen in liquid nitrogen. RNA was extracted with TRIzol (Invitrogen) according to the manufacturer's instructions and the RNeasy kit (Qiagen) with on-column DNase (Qiagen) digestion. RNA was quantified and the quality was checked with a 2100 Bioanalyzer (Agilent).

RNA sequencing analysis

Library preparation was done using the TruSeq RNA Sample Preparation Kit v2 (Illumina). In brief, polyA-containing mRNA molecules were reverse transcribed, double-stranded cDNA was generated and adapters were ligated. After quality control using 2100 Bioanalyzer (Agilent), clusters were generated through amplification using the TruSeq PE Cluster Kit v3-cBot-HS kit (Illumina) followed by sequencing on a Illumina HiSeq2000 with the TruSeq SBS Kit v3-HS (Illumina). Sequencing was performed in Paired-End mode with a read length of 50 bp. The

quality of the raw data was verified with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/, version 0.9.1). Next, quality filtering was performed using FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/, version 0.0.13): reads where globally filtered in which for at least 75% of the reads the quality exceeds Q20 and 3' trimming was performed to remove bases with a quality below Q10. Re-pairing was performed using a custom Perl script. Reads were subsequently mapped to the Arabidopsis reference genome (TAIR10) using GSNAP (Wu and Nacu, 2010) (version 2011-12-28) allowing maximally two mismatches. The concordantly paired reads that uniquely map to the genome were used for quantification on the gene level with htseq-count from the HTSeq.py python package (Anders et al., 2015). The analysis was implemented as a workflow in Galaxy (Goecks et al., 2010).

For the visualization of RNA-seq expression data and correlation analysis, count data was normalized following the normalization pipeline with the trimmed mean of M-values (TMM) algorithm as implemented in the edgeR library from the R software (v.3.0.1) (R Core Team, 2015). Weakly expressed genes were previously filtered out by removing genes that have less than five samples with an expression level lower than 10 counts per million. The normalized count data were then transformed with inverse hyperbolic sine function "asinh" in R software (v.3.0.1) (R Core Team, 2015).

The PCA plot on transformed count data was done in R using 'pca' function.

Sequence extraction and alignment

The different read libraries were quality checked using FastQC v0.9.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapters were trimmed using cutadapt in python v3.1.1 with the options: --adapter 'GATCGGAAGAGCACACGTCTGAACTCCAGTCAC', --overlap 10 and --minimum-length 35 (Martin, 2011). Quality filtering was performed using Fastx v0.0.13 with the options: -Q 33, -q 10 and -p 75 (http://hannonlab.cshl.edu/fastx_toolkit/). Reads were also trimmed using Fastx v0.0.13 with the options: -Q 33, -t 20 and -l 35. Forward and reverse reads were subsequently collapsed into a single file. After preprocessing, the different read libraries were mapped to the TAIR10 genome using gsnap v2013-02-05 with the options --trim-mismatch-score -3, -k 15, -A sam, -B 4, -n 50, -w 15000, -a off, --quality-protocol sanger, --pairmax-rna 15000, -N 1 and -m 5 (Wu and Nacu, 2010). Only uniquely mapping reads were further considered. Next, sorting and deduplication of the read libraries was performed using picard v1.129 (http://broadinstitute.github.io/picard/). GATK v3.3.0 was used for variant calling (Van der Auwera et al., 2013). Analysis was based on recommendations in 'Best practices for RNA-seq' (https://www.broadinstitute.org/gatk/guide/best-practices?bpm=RNAseq). Before variant calling was performed, the different libraries were preprocessed using the tools splitnCigar, haplotypecaller, realignertargetcreator, indelrealigner, baserecalibrator and printreads. In the haplotypecaller step only high quality scores were considered by setting a quality of 50. Next, a

multi-sample variant calling was performed using haplotypecaller. In this step, all samples are analysed together. Variants were filtered using VariantFiltration with the options -window 35, -cluster 3, -filterName FS, -filter "FS > 30.0", -filterName QD and -filter "QD < 2.0". The resulting variants file was split by sample using bcftools (http://github.com/samtools/bcftools). Sequences were extracted for the genes (AT4G14713, AT4G14720, AT1G19270, AT2G01570 and AT4G25420) using the alternative alleles for each sample using the GATK tool Fasta Alternate Reference Maker (Van der Auwera et al., 2013) and based on the CDS coordinates (based on the structural annotation of TAIR10). The reverse complement was generated for genes located at the negative strand and subsequently protein sequences were extracted using custom scripts.

To align extracted sequence, CLC main Workbench 6.0 was used (CLC bio, a QIAGEN Company; http://www.clcbio.com/).

Differential expression analysis

Differential expression analyses of RNA-seq data were conducted with the 'EdgeR' library (v.3.4.2) of the Bioconductor software from the R software (v.3.0.1) (R Core Team, 2015). Filtering and normalization was performed as previously described. In this analysis, we consider transgenic lines of a particular accession as repeats of a single line, otherwise we would not be able to run statistical tests, because we have a single repeat per line. Two statistical tests were performed; the first test for general, mean differential expression between wild types and transgenic lines of these accessions. The second test is for genes differentially expressed between the wild type of an accession and transgenic lines of that accession in at least one accession. The tests were run using the glmLRT function with a contrast adequate to the test. Next, P values from both tests were corrected for multiple testing with FDR using the qvalue function from 'qvalue' package (v.1.36.0) (Storey et al., 2015) for Bioconductor. This results in two q values (corrected p values from both tests) for each gene. The lowest value was assigned as the final q value for a gene. For further analysis, genes were selected based on an FDR value lower than 0.05 and/or fold change threshold between transgenic lines and wild type. The filter on the fold change requires a fold change higher than 1.5 for each transgenic line of an accession in at least one accession.

Enrichment analysis was done in mapman (Ramšak et al., 2014) (http://mapman.mpimp-golm.mpg.de/pageman/) with DE genes filtered for an FDR value lower than 0.05.

Heat maps are generated in Mev (v 4.9) (Howe et al., 2011) for DE gene filtered for an FDR value lower than 0.05 and a 1.5-fold change threshold between transgenic lines and the wild type. Hierarchical clustering was done for both genes and samples with Manhattan distance metrics in Mev (v 4.9) (Howe et al., 2011).

Correlation analysis

Pearson correlation coefficients were calculated with corr.test function in R. The adjusted P values of correlations were calculated with a permutation test. For each analysis, namely, the phenotype-phenotype, expressivity-RNA-seq fold change and RNA-seq fold change-RNA-seq fold change we ran 1000 permutations. Each permutation run involves a permutation of a single phenotype or expressivity, respectively to analysis, and calculation of correlations between permuted vector and the rest of analyzed data. The adjusted P values over all permutation runs per analysis were calculated as a proportion of correlation coefficients correlated in a higher degree than a tested correlation (r) to the number of permuted correlations (n); with a formula (r+1)/(n+1) (North et al., 2002). The significant correlations, FDR<0.05, were visualized in Cytoscape (Cline et al., 2007).

Extended Data

All extended data is listed below. Extended data can be found at the end of this chapter.

Extended Data Figure 6.1 Sequence alignments of the endogenous cDNA and protein sequence of PPD2 in 15 *Arabidopsis* accessions.

Extended Data Table 6.1 Phenotype of *amiPPD* transgenics in 15 Arabidopsis accessions per independent transgenic line.

References

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq – a Python framework to work with high-throughput sequencing data. Bioinformatics **31:** 166-169.

Andriankaja, M., Dhondt, S., De Bodt, S., Vanhaeren, H., Coppens, F., De Milde, L., Mühlenbock, P., Skirycz, A., Gonzalez, N., Beemster, G.T.S., and Inzé, D. (2012). Exit from proliferation during leaf development in *Arabidopsis thaliana*: a not-so-gradual process. Dev. Cell **22:** 64-78.

Bai, Y., Meng, Y., Huang, D., Qi, Y., and Chen, M. (2011). Origin and evolutionary analysis of the plant-specific TIFY transcription factor family. Genomics **98:** 128-136.

Bergmann, D.C., and Sack, F.D. (2007). Stomatal development. Annu. Rev. Plant Biol. **58:** 163-181.

Chandler, C.H., Chari, S., and Dworkin, I. (2013). Does your gene need a background check? How genetic background impacts the analysis of mutations, genes, and evolution. Trends Genet. **29:** 358-366.

Chini, A., Fonseca, S., Chico, J.M., Fernández-Calvo, P., and Solano, R. (2009). The ZIM domain mediates homo- and heteromeric interactions between Arabidopsis JAZ proteins. Plant J. **59:** 77-87.

Claeys, H., Skirycz, A., Maleux, K., and Inzé, D. (2012). DELLA signaling mediates stress-induced cell differentiation in Arabidopsis leaves through modulation of anaphase-promoting complex/cyclosome activity. Plant Physiol. **159:** 739-747.

Cline, M.S., et al. (2007). Integration of biological networks and gene expression data using Cytoscape. Nat. Protoc. **2:** 2366-2382.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:** 735-743.

Cuellar Perez, A., Nagels Durand, A., Vanden Bossche, R., De Clercq, R., Persiau, G., Van Wees, S.C., Pieterse, C.M., Gevaert, K., De Jaeger, G., Goossens, A., and Pauwels, L. (2014). The non-JAZ TIFY protein TIFY8 from Arabidopsis thaliana is a transcriptional repressor. PLoS ONE **9:** e84891.

De Bodt, S., Hollunder, J., Nelissen, H., Meulemeester, N., and Inze, D. (2012). CORNET 2.0: integrating plant coexpression, protein-protein interactions, regulatory interactions, gene associations and functional annotations. New Phytol **195:** 707-720.

De Bodt, S., Carvajal, D., Hollunder, J., Van den Cruyce, J., Movahedi, S., and Inzé, D. (2010). CORNET: a user-friendly tool for data mining and integration. Plant Physiol. **152:** 1167-1179.

Elsner, J., Michalski, M., and Kwiatkowska, D. (2012). Spatiotemporal variation of leaf epidermal cell growth: a quantitative analysis of *Arabidopsis thaliana* wild-type and triple *cyclinD3* mutant plants. Ann. Bot. **109:** 897-910.

Geisler, M., Nadeau, J., and Sack, F.D. (2000). Oriented asymmetric divisions that generate the stomatal spacing pattern in Arabidopsis are disrupted by the *too many mouths* mutation. Plant Cell **12:** 2075-2086.

Goecks, J., Nekrutenko, A., Taylor, J., and The Galaxy Team. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol. **11:** R86.

Gonzalez, N., Vanhaeren, H., and Inzé, D. (2012). Leaf size control: complex coordination of cell division and expansion. Trends Plant Sci. **17:** 332-340.

Gonzalez, N., et al. (2010). Increased leaf size: different means to an end. Plant Physiol. **153:** 1261-1279.

Gonzalez, N., et al. (2015). A Repressor Protein Complex Regulates Leaf Growth in Arabidopsis. Plant Cell **27:** 2273-2287.

Gupta, R., Webster, C.I., and Gray, J.C. (1997). The single-copy gene encoding high-mobility-group protein HMG-I/Y from pea contains a single intron and is expressed in all organs. Plant Mol. Biol. **35:** 987-992.

Horvath, D.P., Chao, W.S., and Anderson, J.V. (2002). Molecular analysis of signals controlling dormancy and growth in underground adventitious buds of leafy spurge. Plant Physiol. **128:** 1439-1446.

Horvath, D.P., Schaffer, R., West, M., and Wisman, E. (2003). Arabidopsis microarrays identify conserved and differentially expressed genes involved in shoot growth and development from distantly related plant species. The Plant Journal **34:** 125-134.

Howe, E.A., Sinha, R., Schlauch, D., and Quackenbush, J. (2011). RNA-Seq analysis in MeV. Bioinformatics **27:** 3209-3210.

Lau, O.S., Davies, K.A., Chang, J., Adrian, J., Rowe, M.H., Ballenger, C.E., and Bergmann, D.C. (2014). Direct roles of SPEECHLESS in the specification of stomatal self-renewing cells. Science **345:** 1605-1609.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal **17:** 10-12.

North, B.V., Curtis, D., and Sham, P.C. (2002). A note on the calculation of empirical *P* values from Monte Carlo procedures. Am. J. Hum. Genet. **71:** 439-441.

Parizotto, E.A., Dunoyer, P., Rahm, N., Himber, C., and Voinnet, O. (2004). In vivo investigation of the transcription, processing, endonucleolytic activity, and functional relevance of the spatial distribution of a plant miRNA. Genes Dev **18:** 2237-2242.

Pauwels, L., et al. (2010). NINJA connects the co-repressor TOPLESS to jasmonate signalling. Nature **464:** 788-791.

R Core Team. (2015). R: a language and environment for statistical computing. Vienna, Austria (http://www.R-project.org/).

Radhakrishnan, I., Pérez-Alvarado, G.C., Parker, D., Dyson, H.J., Montminy, M.R., and Wright, P.E. (1997). Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. Cell **91:** 741-752.

Rae, G.M., Uversky, V.N., David, K., and Wood, M. (2014). DRM1 and DRM2 expression regulation: potential role of splice variants in response to stress and environmental factors in Arabidopsis. Mol Genet Genomics **289:** 317-332.

Ramšak, Ž., Baebler, Š., Rotter, A., Korbar, M., Mozetič, I., Usadel, B., and Gruden, K. (2014). GoMapMan: integration, consolidation and visualization of plant gene annotations within the MapMan ontology. Nucleic Acids Res. **42:** D1167-1175.

Reeves, R. (2001). Molecular biology of HMGA proteins: hubs of nuclear function. Gene **277:** 63-81.

Reeves, R., and Nissen, M.S. (1993). Interaction of high mobility group-I (Y) nonhistone proteins with nucleosome core particles. J Biol Chem **268:** 21137-21146.

Ruyter-Spira, C., de Groof, A., van der Poel, J.J., Herbergs, J., Masabanda, J., Fries, R., and Groenen, M. (1998). The HMGI-C gene is a likely candidate for the autosomal dwarf locus in the chicken. J. Hered. **89:** 295-300.

Shikata, M., Takemura, M., Yokota, A., and Kohchi, T. (2003). Arabidopsis ZIM, a plant-specific GATA factor, can function as a transcriptional activator. Biosci. Biotechnol. Biochem. **67:** 2495-2497.

Shimada, T.L., Shimada, T., and Hara-Nishimura, I. (2010). A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. Plant J. **61:** 519-528.

Shpak, E.D., Berthiaume, C.T., Hill, E.J., and Torii, K.U. (2004). Synergistic interaction of three ERECTA-family receptor-like kinases controls Arabidopsis organ growth and flower development by promoting cell proliferation. Development **131:** 1491-1501.

Storey, J.D., Bass, A.J., Dabney, A., and Robinson, D. (2015). qvalue: Q-value estimation for false discovery rate control. R package version 2.2.0 (http://www.bioconductor.org/packages/release/bioc/html/qvalue.html).

Van der Auwera, G.A., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics **11:** 11 10 11-11 10 33.

Vos, J.W., Pieuchot, L., Evrard, J.-L., Janski, N., Bergdoll, M., de Ronde, D., Perez, L.H., Sardon, T., Vernos, I., and Schmit, A.-C. (2008). The plant TPX2 protein regulates prospindle assembly before nuclear envelope breakdown. Plant Cell **20:** 2783-2797.

White, D.W.R. (2006). *PEAPOD* regulates lamina size and curvature in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **103:** 13238-13243.

Wu, T.D., and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics **26:** 873-881.

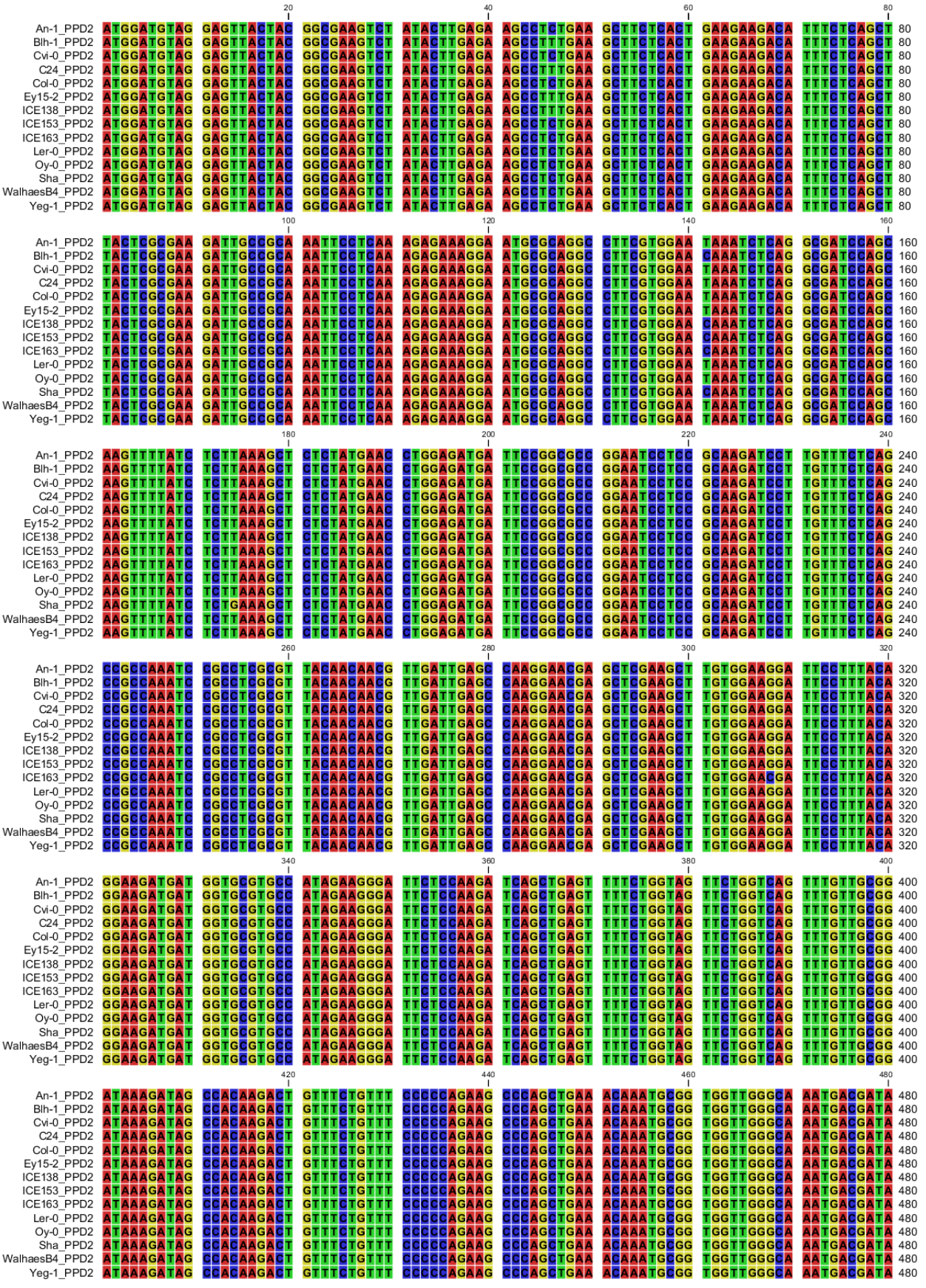
Zhang, W., Wu, Q., Pwee, K.-H., and Kini, R.M. (2003). Interaction of wheat high-mobility-group proteins with four-way-junction DNA and characterization of the structure and expression of HMGA gene. Arch. Biochem. Biophys. **409:** 357-366.

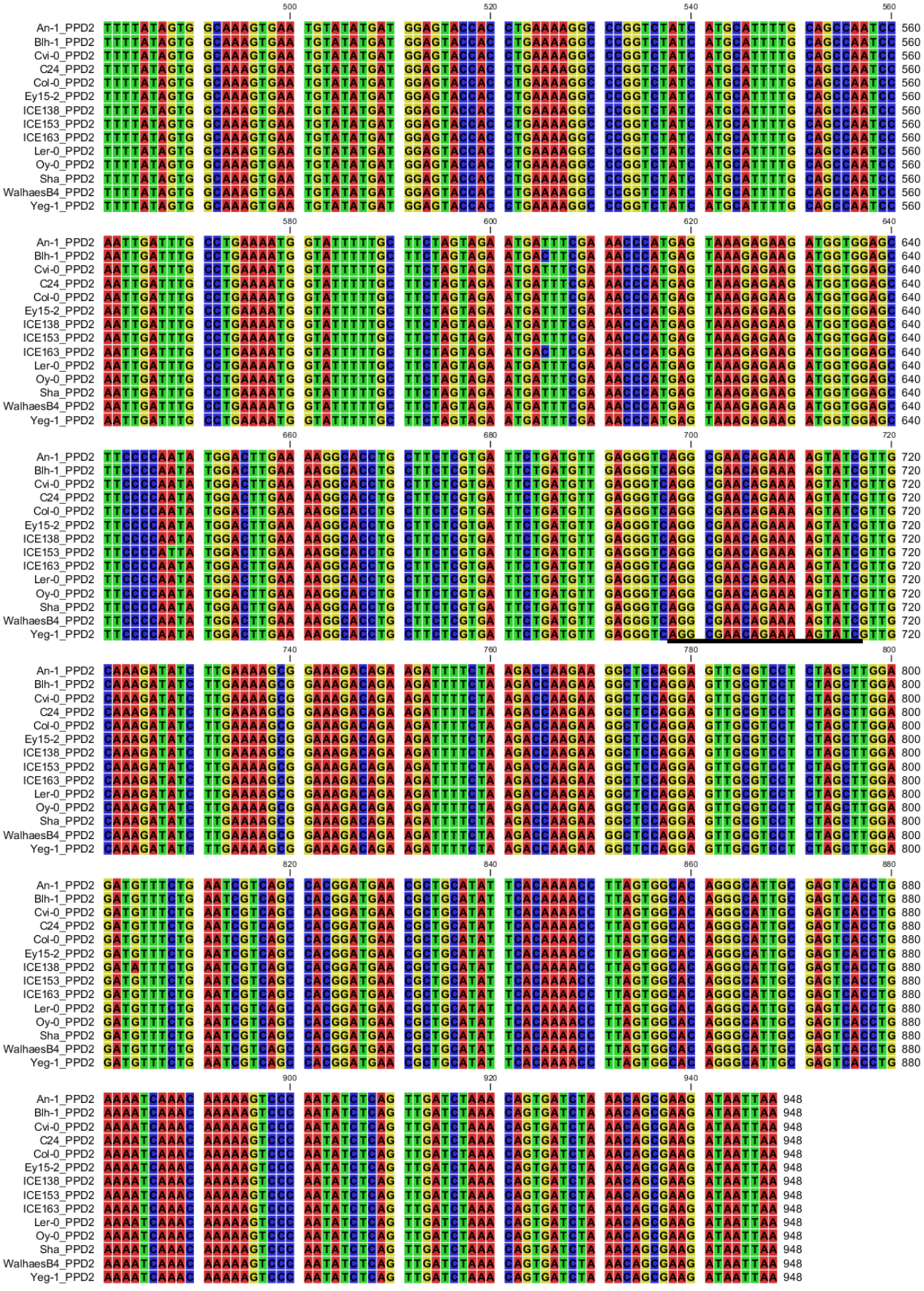
Zhao, J., and Grafi, G. (2000). The high mobility group I/Y protein is hypophosphorylated in endoreduplicating maize endosperm cells and is involved in alleviating histone H1-mediated transcriptional repression. J. Biol. Chem. **275:** 27494-27499.

Zhao, J., Paul, L.K., and Grafi, G. (2009). The maize HMGA protein is localized to the nucleolus and can be acetylated in vitro at its globular domain, and phosphorylation by CDK reduces its binding activity to AT-rich DNA. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms **1789:** 751-757.

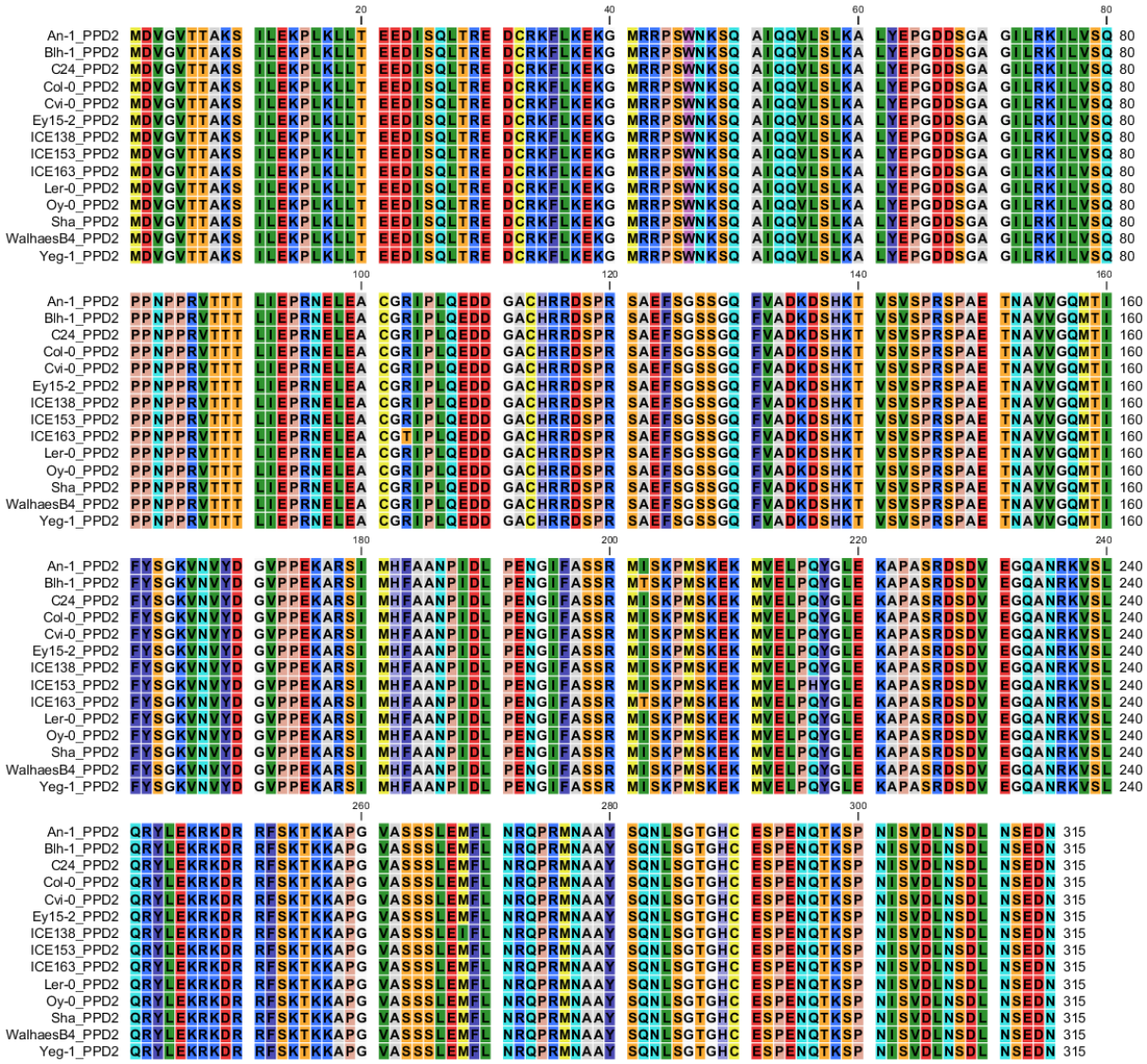
Zhou, X., Benson, K.F., Ashar, H.R., and Chada, K. (1995). Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. Nature **376:** 771-774.

A


An-1_PPD2
Blh-1_PPD2
Cvi-0_PPD2
C24_PPD2
Col-0_PPD2
Ey15-2_PPD2
ICE138_PPD2
ICE153_PPD2
ICE163_PPD2
Ler-0_PPD2
Oy-0_PPD2
Sha_PPD2
WalhaesB4_PPD2
Yeg-1_PPD2


500
520
540
560
An-1_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
Blh-1_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
Cvi-0_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
C24_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
Col-0_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
Ey15-2_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
ICE138_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
ICE153_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
ICE163_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
Ler-0_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
Oy-0_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
Sha_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
WalhaesB4_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
Yeg-1_PPD2 TTTTATAGTG GCAAAGTGAA TGTATATGAT GGAGTACCAC CTGAAAAGGC CCGGTCTATC ATGCATTTTG CAGCCAATCC 560
580
600
620
640
An-1_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGATTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
Blh-1_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGACTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
Cvi-0_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGATTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
C24_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGATTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
Col-0_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGATTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
Ey15-2_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGATTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
ICE138_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGATTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
ICE153_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGATTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
ICE163_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGACTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
Ler-0_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGATTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
Oy-0_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGATTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
Sha_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGATTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
WalhaesB4_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGATTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
Yeg-1_PPD2 AATTGATTTG CCTGAAAATG GTATTTTTGC TTCTAGTAGA ATGATTTCGA AACCCATGAG TAAAGAGAAG ATGGTGGAGC 640
660
680
700
720
An-1_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
Blh-1_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
Cvi-0_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
C24_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
Col-0_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
Ey15-2_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
ICE138_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
ICE153_PPD2 TTCCCCATTA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
ICE163_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
Ler-0_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
Oy-0_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
Sha_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
WalhaesB4_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
Yeg-1_PPD2 TTCCCCAATA TGGACTTGAA AAGGCACCTG CTTCTCGTGA TTCTGATGTT GAGGGTCAGG CGAACAGAAA AGTATCGTTG 720
740
760
780
800
An-1_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
Blh-1_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
Cvi-0_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
C24_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
Col-0_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
Ey15-2_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
ICE138_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
ICE153_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
ICE163_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
Ler-0_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
Oy-0_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
Sha_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
WalhaesB4_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
Yeg-1_PPD2 CAAAGATATC TTGAAAAGCG GAAAGACAGA AGATTTTCTA AGACCAAGAA GGCTCCAGGA GTTGCGTCCT CTAGCTTGGA 800
820
840
860
880
An-1_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
Blh-1_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
Cvi-0_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
C24_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
Col-0_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
Ey15-2_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
ICE138_PPD2 GATATTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
ICE153_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
ICE163_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
Ler-0_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
Oy-0_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
Sha_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
WalhaesB4_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
Yeg-1_PPD2 GATGTTTCTG AATCGTCAGC CACGGATGAA CGCTGCATAT TCACAAAACC TTAGTGGCAC AGGGCATTGC GAGTCACCTG 880
900
920
940
An-1_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
Blh-1_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
Cvi-0_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
C24_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
Col-0_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
Ey15-2_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
ICE138_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
ICE153_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
ICE163_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
Ler-0_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
Oy-0_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
Sha_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
WalhaesB4_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948
Yeg-1_PPD2 AAAATCAAAC AAAAAGTCCC AATATCTCAG TTGATCTAAA CAGTGATCTA AACAGCGAAG ATAATTAA 948

B



Extended Figure 6.1 Sequence alignments of the endogenous cDNA (A) and protein (B) sequence of PPD2 in 15 *Arabidopsis* accessions.

Extended Data Table 6.1 Phenotype of *amiPPD* transgenics in 15 Arabidopsis accessions per independent transgenic line. Heat maps representing, per independent transgenic line, the average percent difference in each leaf area between *amiPPD* transgenics and their corresponding wild type (left) and the estimated expressivity (middle) of *amiPPD*. More than 10 plants were analyzed. Fold changes of expression levels of *PPD2* compared to wild type analyzed by qPCR in T_2 generation are shown on the right panel (L1-L5; independent transgenic lines).

Accession	Lines	cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15	Sel	Ros	PPD2
ICE61	L1	119	134	134	135	147	154	118	102	89	82	98	80	86			139	118	0.45
	L2	103	124	120	127	148	144	115	90	69	64	72	66	66			131	108	0.38
	L3	87	104	105	112	110	113	104	93	81	81	99	95	111			109	98	0.48
	L4	108	116	121	122	136	143	111	94	76	63	84	74	78			126	107	0.45
	L5	117	125	124	130	148	141	129	103	89	83	98	99	107			134	117	0.43
An-1	L1	92	81	107	111	107	111	87									108	130	0.60
	L2	111	116	126	136	127	108	99									129	142	0.66
	L3	151	137	153	161	155	142	119									156	172	0.64
	L4	134	145	143	154	134	116	97									143	164	0.47
	L5	133	129	150	149	146	130	99									148	150	0.26
Cvi-0	L2	125	124	109	108	105	81	76	58	34							107	103	0.62
	L3	109	106	103	99	103	88	76	77	78							102	107	0.57
Ler-0	L1	88	81	95	97	110	101	88	93								100	118	0.25
	L2	105	112	133	136	131	103	72	62								134	115	0.25
	L3	101	109	123	126	135	112	91	95								128	131	0.51
	L4	108	105	138	142	124	104	87	82								135	123	0.63
Blh-1	L1	93	134	134	162	134	128	124	110	83							144	146	0.36
	L2	97	131	146	155	143	134	133	119	114							148	142	0.42
	L3	83	116	127	133	131	129	117	105	93							130	133	0.40
ICE138	L1	128	152	157	157	116	100	83	70	63	58	36	70				143	113	0.43
	L2	102	110	124	130	117	100	84	63	52	42	28	36				124	101	0.61
	L4	99	99	114	102	111	117	98	101	108	82	89	94				109	115	0.58
Ey152	L1	90	127	138	163	139	141	121	100	95	63	64	48	31	29	19	147	98	0.44
	L2	119	131	145	132	125	116	116	97	86	82	75	69	64	70	71	134	100	0.38
	L3	78	79	72	68	64	73	79	76	74	65	63	49	39	55	52	68	68	0.46
	L4	126	131	131	124	107	121	119	107	94	88	80	65	66	68	45	121	104	0.37
	L5	119	129	134	129	121	122	114	93	71	68	55	49	46	53	56	128	91	0.37
C24	L1	94	112	118	120	122	118	108	97	86	62	48	39	37	29		120	89	0.51
	L2	84	101	115	108	111	106	91	85	56	40	32	22	26	17		111	76	0.48
	L3	99	106	107	111	111	103	107	97	85	80	65	56	50	44		110	92	0.43
	L4	64	82	100	102	105	93	79	51	26	16	10	6	5	5		103	57	0.46
	L5	102	104	114	121	120	117	123	115	94	85	67	57	49	46		118	99	0.49
Yeg-1	L2	94	139	130	146	131	132	122	94	80	55	53	39	39			136	111	0.38
	L3	78	115	112	114	100	89	66	45	30	18	13	10	14			109	74	0.50
	L4	55	73	113	106	119	137	124	113	114	92	110	81	101			113	117	0.44
WalhaesB4	L1	97	137	118	123	109	101	96	84	86	66	62	55	49			117	97	0.39
	L4	96	132	125	117	102	104	93	82	71	60	47	45	40			115	93	0.40
Sha	L2	91	114	120	118	113	109	90	73	60	64	41					117	106	0.26
	L3	110	130	138	134	127	124	122	99	82	100	73					133	125	0.29
	L5	108	124	139	134	124	119	96	94	60	80	93					133	122	0.29
Col-0	L1	100	112	115	124	126	110	98	88	81	71	49					122	104	0.35
	L2	123	141	141	143	140	134	119	107	113	89	73					141	126	0.26
	L3	106	126	129	130	121	117	107	101	91	83	83					127	114	0.37
	L4	135	133	140	140	136	128	101	87	83	71	57					138	117	0.38
	L5	118	140	147	151	140	125	103	98	88	76	81					146	121	0.40
ICE153	L1	93	111	125	125	116	106	100	79	72	55	52	46	43			122	99	0.27
	L2	111	119	125	126	121	124	99	99	88	65	77	55	69			124	108	0.33
	L3	92	101	103	106	108	101	99	96	93	80	81	75	59			106	98	0.46
	L4	92	111	113	116	112	112	97	80	62	61	55	48	43			114	97	0.3
ICE163	L1	101	99	98	95	101	103	100	102	103	99	100	99	97	90	87	98	99	0.40
	L4	103	99	98	100	100	100	95	94	95	90	91	89	87	80	76	99	95	0.37
Oy-0	L1	113	122	130	140	133	132	118	117	96	90	76	63	49	54	30	134	109	0.30
	L2	132	122	129	132	134	129	127	117	105	113	90	94	84	91	63	132	116	0.40
	L3	121	117	125	128	134	126	122	120	96	94	86	69	68	61	51	129	110	0.26
	L4	98	95	108	111	116	113	105	106	81	84	72	60	49	49	35	112	95	0.27
	L5	97	107	118	129	126	130	100	93	81	61	66	52	42	57	40	124	97	0.48

Chapter 7

Comparative analysis of natural variation responses to overexpression of *GA20ox1*, dominant negative form of *DA1*, and *amiPPD* in Arabidopsis

Youn-Jeong Nam[1,2], Dorota Herman[1,2], Eunyoung Chae[3], Frederik Coppens[1,2], Bram Slabbinck[1,2], Veronique Storme[1,2], Twiggy Van Daele[1,2], Detlef Weigel[3], Dirk Inzé[1,2*] & Nathalie Gonzalez[1,2*]

[1]Department of Plant Systems Biology, VIB B-9052 Gent, Belgium, [2]Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Gent, Belgium, [3]Department of Molecular Biology, Max Planck Institute for Developmental Biology, Tübingen, Germany
* These authors contributed equally to this work.

Contributions: Y.J.N. was the main author of this work. Y.J.N., E.C, and T.V.D. conducted experimental work. D.H., F.C., B.S., V.S., and Y.J.N. analysed the data. D.I., N.G., and D.W. supervised the research. D.I. and N.G. contributed to the writing of this chapter.

Manuscript is in preparation and will be submitted for publication after minor changes to text.

Introduction

The relationship between a phenotype and a particular mutation cannot only be explained by the presence of a single alteration in the DNA. When transgenes or mutations are introduced into different genetic backgrounds, various sequence variations in the genome might influence the effect of the gene perturbed (Chandler et al., 2013; Paaby and Rockman, 2014). This phenomenon can be also described as epistasis since the effect of a gene is dependent on interactions with gene variants from different genetic backgrounds. Although such epistatic interactions have been observed in several organisms (Chandler et al., 2013), the knowledge about how genetic background influences the phenotypic output of a mutation in plants is poorly understood. For quantitative traits such as leaf growth in plants, which are regulated by a complex network of genes, it would be interesting to identify the causality of these interactions although probably complex due to the amount of genes involved.

To understand natural variation response to genetic perturbation affecting leaf size in Arabidopsis, we chose to modify the expression of three genes known to positively regulate leaf growth in Col-0 background when their expression is altered: overexpression of *GA20ox1* (*GA20ox1*^{OE})(Gonzalez et al., 2010), overexpression of an artificial microRNA targeting *PPD2 (amiPPD)* (Gonzalez et al., 2015), and overexpression of a dominant negative form of *DA1 (DN-DA1*^{OE}*)* (Li et al., 2008). The three genes encode proteins which have different biological functions: GA20ox1 is a rate-limiting gibberellin biosynthesis enzyme, PPD2 is a transcription regulator, and DA1 is an ubiquitin receptor (see chapter 2 for details). The transgenes were introduced in seventeen Arabidopsis natural accessions from around the world including Col-0 (Figure 3.1 and Table 4.1). Leaf size of all transgenic lines was measured with leaf series and transcriptome analysis was conducted to examine changes at molecular level (for *GA20ox1*^{OE} transgenic lines, see chapter 5 and for *amiPPD* transgenic lines, see chapter 6).

In this research chapter, we firstly describe the effect of *DN-DA1*^{OE} transgene in 17 accessions and afterwards we conducted a comparative analysis of the effects of the three different transgenes in the different genetic backgrounds.

Results

Phenotypic analysis of plants overexpressing a dominant negative form of *DA1* in 17 accessions

DA1 encodes an ubiquitin receptor controlling final seed and organ size in Col-0 (Li et al., 2008) and is expressed in the whole plant. A point mutation in *DA1* gene sequence, in the *da1-1* mutant, leads to the production of larger round leaves, larger flowers and seeds, and thicker siliques. This point mutation leads to a nonsynonymous substitution at a conserved amino acid changing arginine to lysine at position 358 and creating a dominant negative form of DA1. Transgenic plants overexpressing a construct harboring the DA1^{R358K} allele phenocopy the *da1-1*

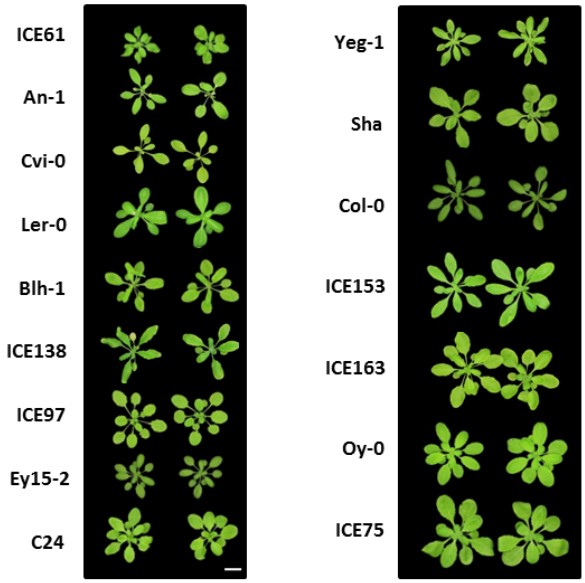
mutant. The improved growth phenotype in this dominant negative *DA1* overexpressing (*DN-DA1^{OE}*) line as well as in the *da1-1* mutant was observed in Col-0 background (Li et al., 2008).

In order to analyze the transferability of the positive effect on leaf growth of the mutation in *DA1* in different genetic backgrounds, T_3 homozygote *DN-DA1^{OE}* transgenics lines were generated in 17 Arabidopsis accessions. For each accession, from 1 to 5 independent transformants showing "the highest overexpression" still with a range of expression (at least more than 2 fold change compared to wild type) and with single locus insertion were selected (Extended Data Table 7.1) and homozygous plants were grown in soil for 25 days (see Methods). Representative pictures of the *DN-DA1^{OE}* transgenic plants in comparison to their respective control accessions (no pictures of WalhaesB4) are shown in Figure 7.1A. Most of the transgenic lines have rounder leaves than their corresponding wild type except the transgenics in Yeg-1 that do not seem to show this phenotype (Figure 7.1A). To quantify leaf size in detail, individual leaf area of all transgenic lines was measured at 25 DAS by making leaf series (Extended Data Table 7.1**)**. In general, most transgenic lines showed consistency of the effect of a transgene within an accession. However, there were some transgenic lines showing different effects on leaf size alteration for instance, line 1 of An-1, line 2 of Ey152, and two transgenic lines of ICE153 (Extended Data Table 7.1**)**. The average effect of independent transgenic lines per accession of the *DN-DA1^{OE}* was estimated (Figure 7.1B). We observed an increased leaf area in most of the transgenic lines and in seven accessions the increase was significant and mainly in older leaves (first leaf pair, leaf 3, and leaf 4). But in eleven accessions a significant decreased leaf area in younger leaves was found (the leaf position was depending on the accession). The transgenic lines of Yeg-1 had leaves showing either an increased or no change in area but did not produce leaves with decreased size compared to wild type. Therefore, most of the accessions show an increase in the area of the older leaves and a decrease in the younger leaves at the timepoint analysed.

To quantify how much *DN-DA1^{OE}* is affecting positively leaf growth in the 17 accessions, we estimated the expressivity of the transgene. Two different methods were applied to calculate this expressivity. First, because not all leaves are positively affected in a transgenic line, the first leaf pair, leaf 3 and 4 (leaves showing the largest increased area compared to wild type in most of the accessions) were chosen for the calculation of the so called "selective leaf" expressivity (Sel). The "selective leaf" expressivity was determined by averaging the percentage to the wild type of these three leaves (Figure 7.1C). By using this method of calculation, we found that the overexpression of DN-DA1 causes a clear positive effect on leaf area in all accessions as high as 141% in Blh-1 whereas Ey15-2 and Col-0 only had a 5% increase compared to wild type over these three leaves. Second, the expressivity at rosette level (Ros) was estimated by calculating the percentages to wild type of the total rosette area per accession (Figure 7.1C). At rosette level, more variable effects of *DN-DA1^{OE}* were observed. In 10 accessions, *DN-DA1^{OE}* transgenics have larger rosette area (more than 5% increase compared to wild type). However, for 7 accessions, ICE61, Ey15-2, C24, Sha, ICE153, ICE163 and ICE75 although a positive 'selective leaf' expressivity was found either a negative outcome or no effect was observed at rosette level.

In conclusion, expression of the *DN-DA1* construct in different accessions causes relatively high positive effects on the area of older leaves but the range of increase in size varies in function of the background.

A



C

Accession	Sel	Ros
ICE61	115	97
An-1	113	109
Cvi-0	120	108
Ler-0	130	120
Blh-1	141	112
ICE138	130	120
ICE97	130	115
Ey15-2	105	84
C24	121	71
Yeg-1	114	112
WalhaesB4	128	110
Sha	131	105
Col-0	105	88
ICE153	106	103
ICE163	121	94
Oy-0	134	116
ICE75	119	93

B

Accession	cot	L12	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15
ICE61	100	117	110	118	111	117	101	85	76	60	59	64	55	72	
An-1	89	114	109	113	107	108	87								
Cvi-0	99	133	110	113	107	94	84	85	**60**						
Ler-0	101	**134**	**121**	**133**	**115**	99	**76**	**59**							
Blh-1	99	**145**	**134**	**141**	**119**	111	92	**74**	**68**	**51**	**46**				
ICE138	118	124	128	127	133	126	98	111	106	93	103	101			
ICE97	96	131	125	133	132	126	115	122	101	92	94	83	92	**71**	76
Ey15-2	77	106	102	107	96	95	82	75	**65**	**54**	**50**	**47**			
C24	90	116	119	124	116	86	72	**48**	**33**	**21**	**16**	**11**	**11**		
Yeg-1	96	**116**	**116**	109	108	97	102	107	106						
WalhaesB4	**74**	**132**	121	**123**	110	118	113	104	102	94	97	97	100	96	
Sha	92	124	**135**	**134**	111	103	80	**56**	**49**	**36**	**43**				
Col-0	95	115	100	100	98	94	82	**75**	**70**	**65**	**55**	**52**	**50**		
ICE153	98	95	101	105	104	112	98	104	99	**79**	**75**	**61**	**70**	**64**	
ICE163	101	**125**	116	**121**	107	114	100	104	95	**68**	75	54	54	57	51
Oy-0	112	**143**	128	**130**	126	125	116	108	110	96	86	83	82	98	96
ICE75	131	131	111	116	103	98	88	82	75	**66**	**66**	**60**	**57**	**57**	**56**
				60<	60-80	80-90	90-95	95-105	105-110	110-120	120-140	>140			

Figure 7.1 Phenotype of transgenics overexpressing a dominant negative form of *DA1* in 17 Arabidopsis accessions. (A) Image of 25-days-old rosette (except Ey15-2, 24-days-old rosette) of representative dominant negative *DA1* overexpressing *(DN-DA1^{OE})* plants and their corresponding wild types. Scale bar: 2 cm. (B) Heat map representing, per accession, the average percent difference in leaf area between *DN-DA1^{OE}* transgenics and their corresponding wild type. Bold and underline: p-value < 0.05. (C) Heat map showing the estimated expressivity (Sel: 'selective leaf' expressivity, Ros: 'rosette' expressivity, see Methods) of *DN-DA1^{OE}*.

Cellular analysis in transgenic lines overexpressing the dominant negative form of *DA1*

In Col-0, the increased organ size in *da1-1* mutant and *DN-DA1^{OE}* transgenic lines results from a prolonged period of cell proliferation leading to a significant increase in cell number (Li et al., 2008). We tested whether different genetic backgrounds show distinct alteration at cellular level upon *DN-DA1*-overexpression. Five transgenic lines from An-1, Col-0, ICE163, Oy-0 and Sha (one representative transgenic line per accession) were selected based on their leaf phenotype

(showing variable 'selective leaf' expressivity, Figure 7.1B and Figure 7.1C). Since *DN-DA1^{OE}* positively affects older leaves, the third leaf from plants grown for 25 days was harvested for the cellular analysis (Figure 7.2). Increased leaf 3 area was observed in all transgenic lines. However, the changes in leaf 3 area were not the same as the average of different independent lines since we chose only one transgenic line per accession. In An-1, ICE163, and Sha the *DN-DA1^{OE}* transgenic plants showed a more than 40% increase in leaf 3 area while Col-0 and Oy-0 showed a less pronounced increase in leaf size (25% for Col-0 and 27% for Oy-0). In all transgenic lines, an increased pavement cell number was found. A similar trend between the effect on leaf 3 area and the increased pavement cell number was observed showing that the main cause of the increase leaf size was a change in cell number. However, we also observed effects in pavement cell area that differed between the accessions. An increase in pavement cell area (19%) was observed in the transgenic line of Sha. Interestingly, in the transgenic line of ICE163, smaller pavement cells (-22%) were found probably leading to a less pronounced increase in leaf 3 area (40%) compared to the other accessions since pavement cell number was highly increased (66%).

In conclusion, increased leaf 3 area in the *DN-DA1^{OE}* transgenic lines was mainly due to an increased pavement cell number. Also depending on the genetic background the degree of the effect of *DN-DA1^{OE}* on leaf growth at cellular level is different.

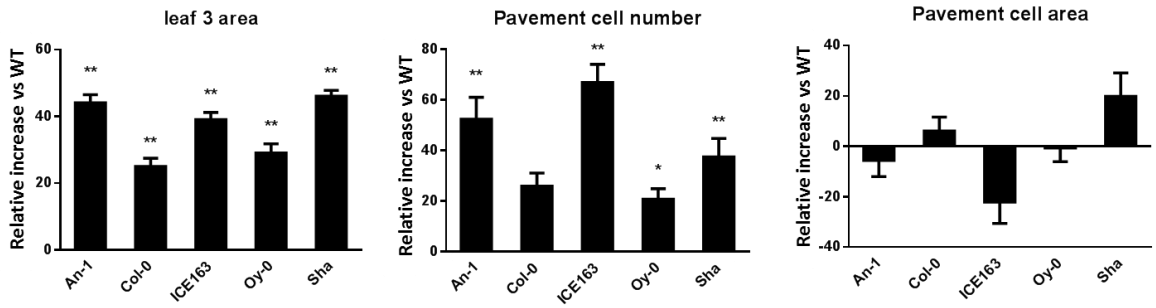


Figure 7.2 Cellular analysis of *DN-DA1^{OE}* transgenic plants in 5 accessions. 5 accessions (An-1, Col-0, ICE163, Oy-0 and Sha) were selected based on the level of expressivity of the transgene. One transgenic line per accession was chosen and grown in soil for 25 days. Three leaves 3 were analyzed per transgenic and per repeat in three biological repeats. Each graph shows the percentage to wild type of leaf 3 area, pavement cell number, and pavement cell area. Values are percentage averages with standard error. (ANOVA ;* adj-P < 0.05, ** adj-P < 0.01)

Comparative analysis of the effects of three transgenes on leaf growth in natural variants

Since we have analyzed the growth response to three different transgenes (*GA20ox1^{OE}, amiPPD* and *DN-DA1^{OE}*) in 17 accessions, a meta-analysis was performed to compare transgene effect in different genetic backgrounds. We compiled all size measurement data of the different transgenic lines for the three transgenes in the 17 accessions for comparative analysis based on the leaf series data and the calculated expressivity (Figure 7.1, Figure 5.2, and Figure 6.1). The results are shown in Figure 7.3

A

	GA20ox1-OE		*amiPPD*		*DN-DA1-OE*	
Accessions	**Sel**	**Ros**	**Sel**	**Ros**	**Sel**	**Ros**
ICE61	133	127	128	110	115	97
An-1	72	53	137	152	113	109
Cvi-0	207	113	104	105	120	108
Ler-0	80	72	124	122	130	120
Blh-1	94	78	141	140	141	112
ICE138	138	124	125	110	130	120
ICE97	367	139	\	\	130	115
Ey15-2	118	95	120	92	105	84
C24	55	80	112	83	121	71
Yeg-1	139	111	119	101	114	112
WalhaesB4	108	96	116	95	128	110
Sha	147	112	128	118	131	105
Col-0	137	104	135	116	105	88
ICE153	114	111	116	101	106	103
ICE163	131	119	99	97	121	94
Oy-0	182	125	126	105	134	116
ICE75	233	133	\	\	119	93

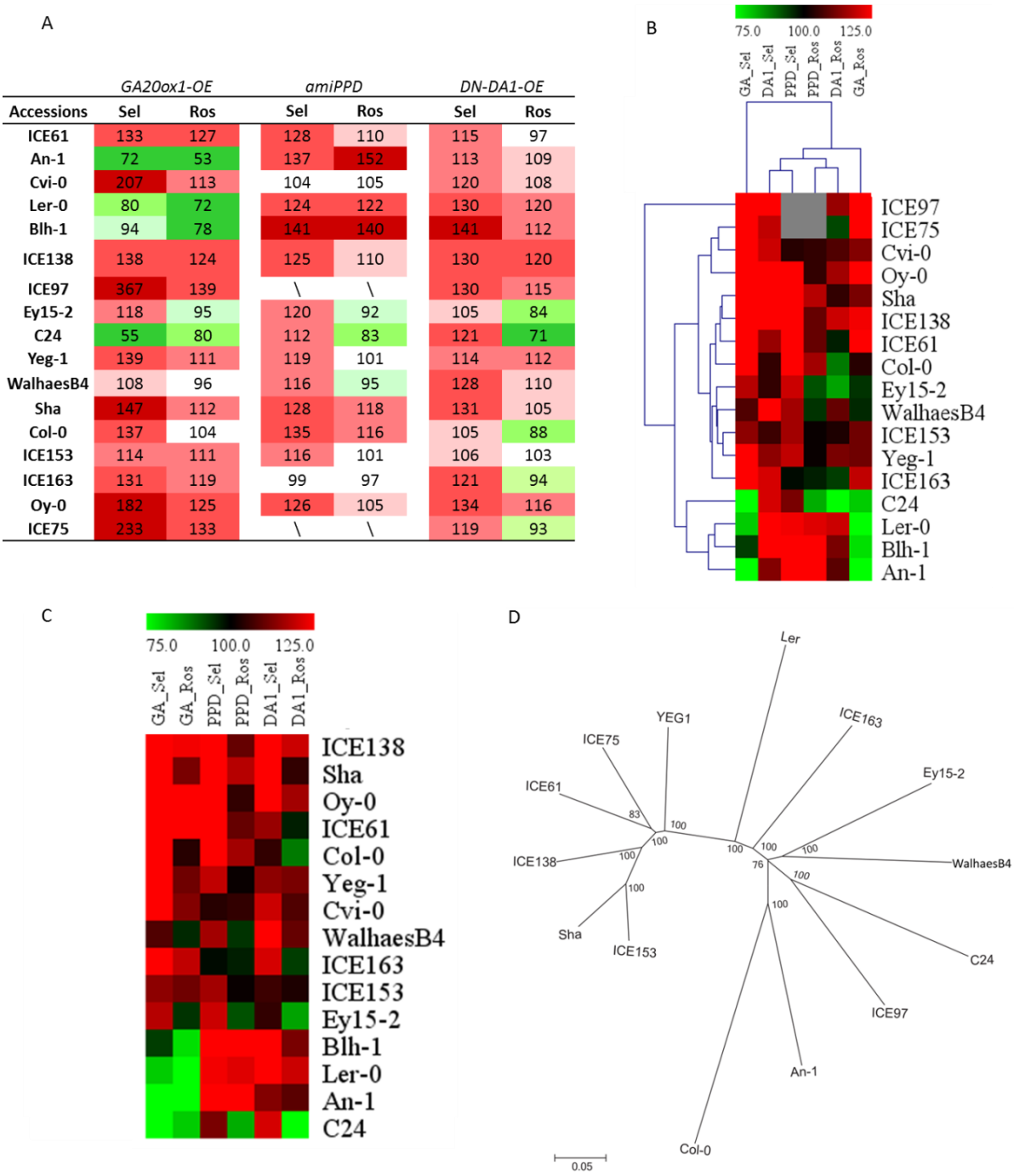


Figure 7.3 Comparison of the expressivity of the three transgenes in 17 accessions. (A) Heat map showing the estimated expressivity (Sel: selective leaf, Ros: rosette, see Material and Methods) of *GA20ox1*^{OE}, *amiPPD,* and *DN-DA1*^{OE}. (B and C) Cluster analysis of expressivity values for the three transgenes in different accessions. Hierarchical clustering was done for both expressivity values and accessions (B) and only for accessions (C) with Manhattan distance metrics. Red color represents high expressivity and green color represents low expressivity in comparison to the wild types. (D) Phylogenetic tree of 14 Arabidopsis accessions used in this study. The phylogenetic tree was constructed using the neighbor-joining method of the MEGA 5 software. The scale bar at bottom represents genetic distance.

In general, the three transgenes had positive effects in most accessions although the position of the leaves affected differed in function of the transgene. Younger leaves were positively affected by overexpression of *GA20ox1* while the *amiPPD* and *DN-DA1^{OE}* constructs positively affected older leaves. Due to the effect of the transgenes on only specific leaves of the rosette, "selective leaf" expressivity values were higher than 'rosettes' expressivity in all transgenic lines. The number of accessions with high "selective leaf" expressivity value is higher in *amiPPD* and *DN-DA1^{OE}* transgenic lines than in *GA20ox1^{OE}* transgenic plants. It seems that the effect of *GA20ox1^{OE}* in the 17 accessions was different from the effect of the two other transgenes. For instance, the positive effect of *GA20ox1^{OE}* was observed at different leaf positions depending on the accession although mostly in younger leaves. In contrast, the two other transgenes showed a somewhat more consistent positive effect on older leaves and as well as negative effect on younger leaves. A cluster analysis using the expressivity values for the three transgenes in the different accessions showed that expressivity of *DN-DA1^{OE}* and *amiPPD* were close to each other while expressivity of *GA20ox1^{OE}* was relatively distant (Figure 7.3B). This analysis also showed that the 'selective leaf' and' rosette' expressivity values were clustered to each other.

Within the same accession, the effects of the different transgenes were also variable (Figure 7.3A). For example, in An-1 transgenic lines, 'rosette' expressivity of *GA20ox1^{OE}* was low (47% decrease) while expressivity of *amiPPD* was high (52% increase) and an intermediate outcome was obtained for *DN-DA1^{OE}* (9% increase). In the case of ICE163, increased rosette area was obtained for *GA20ox1^{OE}* (19% increase), almost no change in size was found in the *amiPPD* transgenic lines and the expressivity of *DN-DA1^{OE}* was low (6% decrease).

Next, in order to analyse in more details how much the effects of the transgenes are influenced by the genetic background, the 17 accessions were clustered based on the expressivity values of the three transgenes (Figure 7.3C). ICE97, for which *amiPPD* lines were not obtained, showed the highest expressivity values for both *GA20ox1^{OE}* and *DN-DA1 ^{OE}* (Figure 7.3B). Among the accessions for which transgenic plants were obtained for the three transgenes, Oy-0 had the highest average expressivity while C24 had the lowest (Figure 7.3C). Among the 17 accessions, in general, ICE138 showed the highest increase in size with both 'rosette' and 'selective leaf' being at least of 10% for the three transgenes. In contrast, ICE153 was the least affected accession since no change was observed in rosette area in both *amiPPD* and *DN-DA1^{OE}* transgenic lines and a maximum of 16% was found for 'selective leaf' of *amiPPD*. From the phylogenetic tree (Figure 7.3D), the two accessions, ICE97 and C24, that showed relatively high and the lowest expressivity values, respectively, are fairly close. Besides, ICE153 and WalhaesB4 which are the least affected by the three transgenes showed genetically long distances within the different accessions. The different responses to the transgenes observed between the accessions seem therefore to not be related to the genetic distance. Also we did not find a clear correlation between geographic distribution of accessions and expressivity

To examine if the accession-specific expressivity could result from differences in sequence of GA20ox1, PPD and DA1 between the accessions and Col-0, the cDNA and protein sequences of the 3 genes were compared (Extended Figure 5.5, Extended Figure 6.1, and Figure 7.4). Some DNA

sequences differences were found in *GA20ox1* and *DA1* (Extended Figure 5.5A and Figure 7.4A), but, in most of the cases these differences led to synonymous substitutions (Extended Figure 5.5B, Figure 7.4B and Figure 7.4C) except for one amino acid insertion (glutamine (Q)) at the conserved position 310 in DA1 of ICE153 (Figure 7.4B) (Li et al., 2008). One of transgenic lines of ICE153 showed larger younger leaves area with round shape (Figure 7.1A)) suggesting that this difference in sequence might not affect the activity of the transgene in the transgenic lines. In general, the sequences of PPD2 proteins showed relatively higher diversity than GA20ox1 and DA1 except for the location of the *amiPPD* target site (Figure 7.4C and Extended Figure 6.1).

In conclusion, the expressivity of the three transgenes in the 17 accessions was dependent on which combination is made between a transgene and a genetic background suggesting the occurrence of complex epistatic interactions. However, the expressivity diversity do not seem to be related to the genetic distance between the accessions neither to the variation in the sequence of the genes studied.

Correlation analysis between wild type phenotypes and expressivity of the transgenes

A correlation analysis was conducted (Figure 7.5) to examine whether the effect of the transgenes, previously defined as 'selective leaf' expressivity and 'rosette' expressivity (Figure 7.3A) could be explained by the phenotype (leaf size-related parameters) of the wild-type accessions (Figure 5.1A).

For *DN-DA1*^{OE} only a negative correlation with leaf number was found (Figure 7.5). A similar negative correlation with leaf number was also found for *amiPPD*. For *amiPPD* transgene, only negative correlations were found between 'rosette' and/or 'selective leaf' expressivity and biomass or rosette area of wild type. In contrast, many positive correlations were found for 'rosette' expressivity of *GA20ox1*^{OE} since biomass, leaf number, rosette area, leaf 1 and 2 area, pavement cell area, and endoreduplication index (EI) were positively correlated with this expressivity corresponding to the percentage increase in total rosette area. 'Selective leaf' expressivity of *GA20ox1*^{OE} was negatively correlated with vascular complexity, vascular density, and cell circularity and similar negative correlations were found with 'rosette' expressivity. Furthermore, the two expressivity values were highly correlated between each other for *amiPPD* and *GA20ox1*^{OE} transgenes while for *DN-DA1*^{OE} a relatively low positive correlation was found (Figure 7.5).

In conclusion, transgenes have different effects in different accessions and, if any, correlation between wild type phenotypes and expressivity values appear to be transgene dependent.

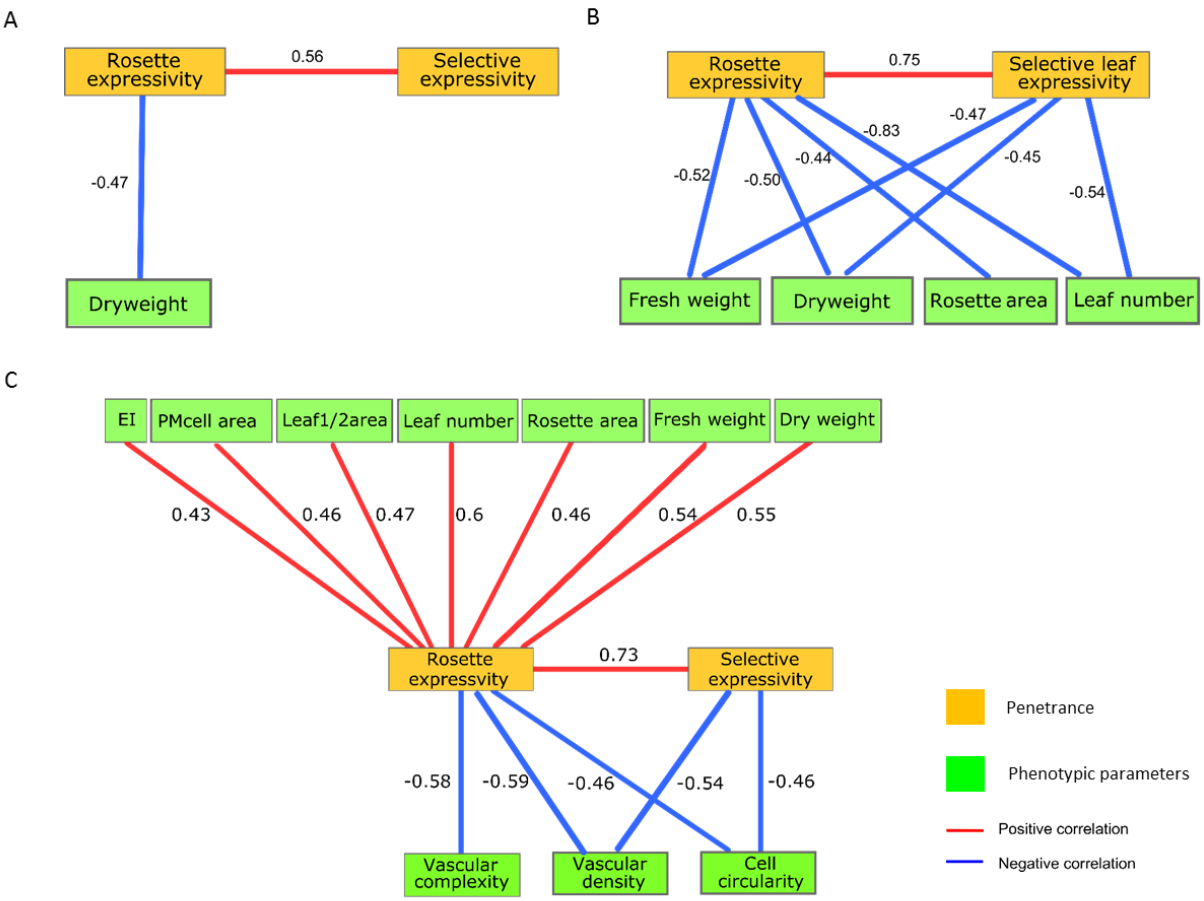


Figure 7.5. Correlation analysis between the phenotypic data for the 3 transgenes. Correlation between phenotypic parameters of wild types and expressivity values for the transgenic lines of *DN-DA1*^{OE} (A), *amiPPD* (B), and *GA20ox1*^{OE} (C). Node color, green; growth-related parameters, Orange; Expressivity. The red and blue edges show positive and negative correlations between parameters, respectively (adj-P value < 0.05). Correlation coefficient values are shown corresponding edges.

Correlation analysis between basal transcript levels in wild type and expressivity

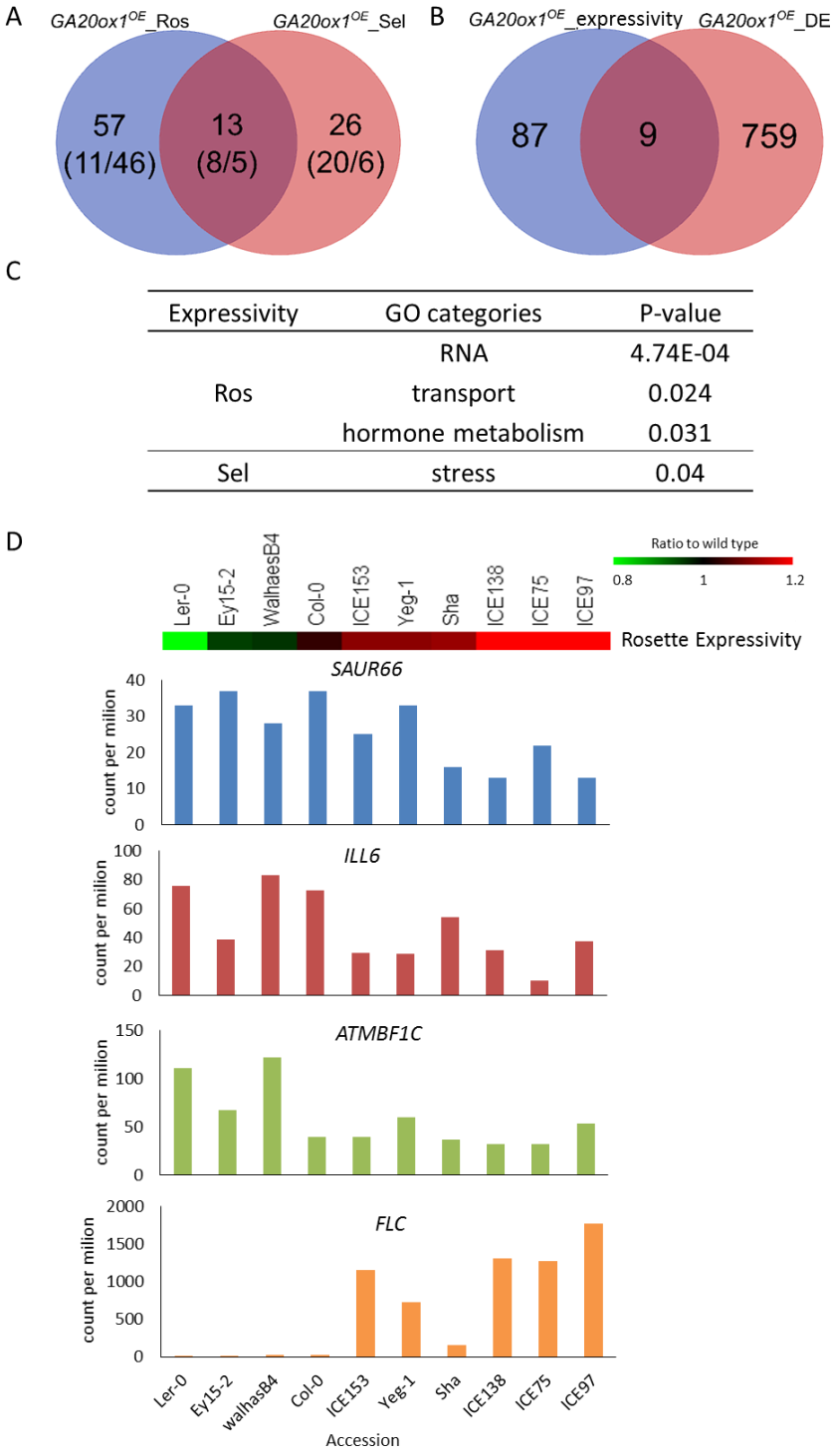
The accurate prediction of background-dependent phenotypic effects of specific mutations has been examined by researchers since it is important for personalized medicine or targeted crop improvement. In C. *elegans*, the effect of genetic background on the severity of mutant phenotypes could be predicted based on variation in the basal expression levels of the target gene in wild type (Vu et al., 2015). By using RNA interference line, loss-of-function phenotypes of 1400 genes were compared in two different genetic backgrounds. This study showed that lower expressed targeted genes leads to more severe phenotype. To check if basal transcript levels of all expressed genes in wild type could be related to the expressivity of the transgenes, we conducted a correlation analysis between transcript expression in wild-type accessions obtained from the RNA sequencing and the expressivity values of *GA20ox1*^{OE} (Figure 7.6 and Table 7.1, can be found at the end of this chapter) and *amiPPD* (Figure 7.7 and Table 7.2, can be found at the end of this chapter). Since we did not obtain transcript expression data from

DN-DA1^{OE}, the correlation analysis was not conducted for this transgenic plants and their corresponding wild types.

For *GA20ox1*^{OE}, the expression levels of a total of 105 genes were found to be correlated with the two expressivity values of *GA20ox1*^{OE} (Figure 7.6A). Almost two times more genes were found for 'rosette' expressivity compared to 'selective leaf' expressivity and 13 genes were overlapping between 'rosette' and 'selective leaf' expressivity. For a majority of genes, expression levels were negatively correlating with 'rosette' expressivity while for numerous genes expression levels were correlating positively with 'selective leaf' expressivity. Interestingly, the genes which expression levels correlate with 'rosette' expressivity belong to similar GO (gene ontology) categories as the DE genes upon *GA20ox1* overexpression (Extended data Table 5.3 and Figure 7.6C). However, most of these genes for which expression in wild type was correlating with expressivity values were not found to be differentially expressed in *GA20ox1* ^{OE} transgenic lines (Figure 7.6B). Four genes involved in hormone metabolism showed negative correlation with 'rosette' expressivity (Figure 7.6D and Table 7.1). Two auxin-related genes encoding SAUR66 and IAA-LEUCINE RESISTANT-LIKE GENE 6 (ILL6), one ethylene-related gene encoding MULTIPROTEIN BRIDGING FACTOR 1C (MBF1C) and one SA-related gene encoding a member of S-ADENOSYL-L-METHYL-DEPENDENT METHYLTRANSFERASES SUPERFAMILY PROTEIN were found. One gene having its expression in wild type positively correlating with 'rosette' expressivity is involved in transcription regulation and encodes the transcription factor *FLOWERING LOCUS C* (*FLC*) that functions as a repressor of floral transition (Deng et al., 2011) (Figure 7.6D and Table 7.1). Several genes related to leaf growth were found to have negative correlations with 'rosette' expressivity. For example, *CAPRICE-LIKE MYB 3* (*CPL3*) has effects not only on flowering development but also on trichome development and epidermal cell size through the regulation of endoreduplication (Tominaga et al., 2008). NITRATE TRANSPORTER 1.1 is involved in redistributing nitrate into developing leaves which is a critical step for optimal plant growth (Hsu and Tsay, 2013). *SHORT INTERNODES/STYLISH* (*SHI*/*STY*) *FAMILY* genes are known to have a function in leaf vein development by regulating local auxin biosynthesis (Baylis et al., 2013).

The expression of a total of 152 genes was found to be correlating with the expressivity values of *amiPPD* (78 genes with 'rosette' expressivity and 85 genes with 'selective leaf' expressivity) (Figure 7.7A). We found that MapMan categories of genes correlating with 'rosette' expressivity have some similarities with GO categories of the DE genes in *amiPPD* transgenic lines but the genes were not the same (Figure 3.5, Figure 7.7B and Figure 7.7C). The expression levels of several genes involved in transport, signalling, and RNA had correlation with 'rosette' expressivity (Table 7.2 and Figure 7.7D). We compared these genes for which expression in wild type is correlating with expressivity with a list of direct targets of PPD2 identified in Col-0 background (Gonzalez et al., 2015). Five of these genes were previously identified as PPD2 direct targets (Table 7.2 and Figure 7.7D). Four genes are positively correlated with expressivity while one gene has negative correlation.

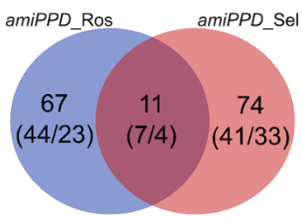
In conclusion, through the correlation analysis between transcript levels in wild type and expressivity of *GA20ox1*^{OE} and *amiPPD* transgenes, genes expressed in wild type involved in similar pathways as DE genes in the transgenic lines were found to correlate with expressivity of the transgenes. However, no correlation was found with basal expression of *PPD2* and *GA20ox1*.



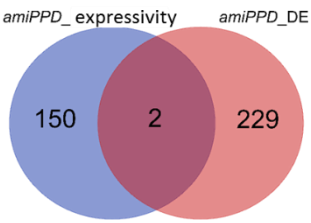
Expressivity	GO categories	P-value
Ros	RNA	4.74E-04
	transport	0.024
	hormone metabolism	0.031
Sel	stress	0.04

Figure 7.6. Correlation analysis between basal expression levels of genes in 10 wild type accessions and expressivity of *GA20ox1*^{OE}. (A) Venn diagram showing the number of genes having their basal expression levels in wild-type correlating with rosette and/or 'selective leaf' expressivity of *GA20ox1*^{OE}. Numbers in bracket correspond to the number of genes having their expression levels showing either a positive correlation (first number) or a negative correlation (second number). (B) Venn diagram representing the number of genes overlapping between genes having their expression correlating with 'rosette' and/or 'selective leaf' expressivity of *GA20ox1*^{OE} and DE genes for *GA20ox1*^{OE}. (C) Over-represented MapMan categories for genes having their expression correlating with 'rosette' and/or 'selective leaf' expressivity of *GA20ox1*^{OE}. (D) Expression level of genes correlating with 'rosette' expressivity of *GA20ox1*^{OE}. Values are count per million from RNA sequencing data. A heat map of 'rosette' expressivity of 10 wild type accessions is shown at the top.

A



B



C

Expressivity	GO categories	P-value
Ros	DNA	3.95E-03
	protein	0.015
	transport	0.022
	signalling	0.025
	fermentation	0.032
	RNA	0.045
Sel	DNA	2.17E-03
	lipid metabolism	0.019
	amino acid metabolism	0.024
	S-assimilation	0.032
	fermentation	0.034

D

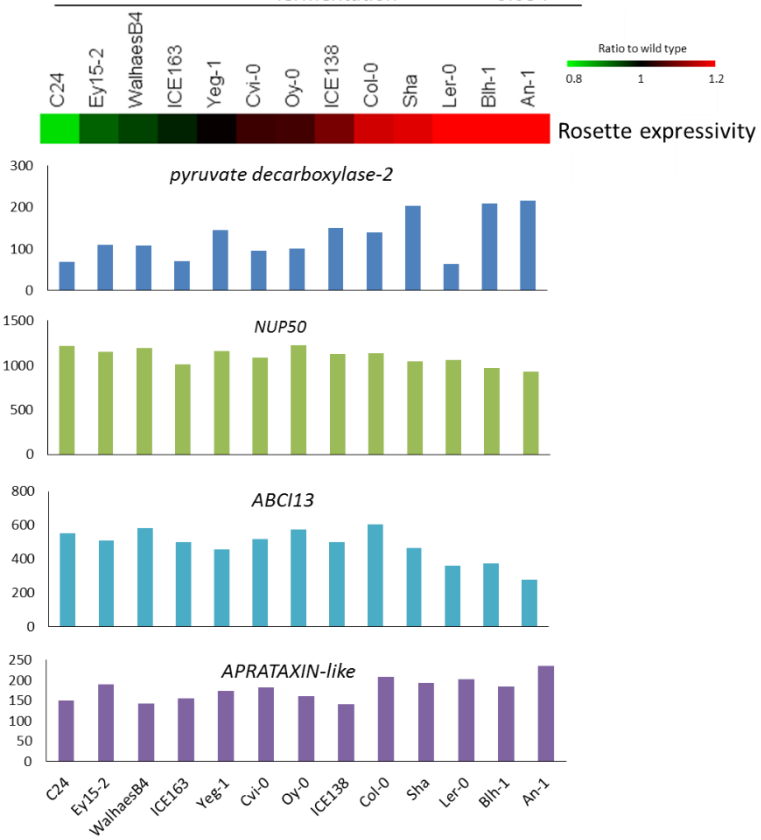


Figure 7.7 Correlation analysis between basal expression levels of genes in 13 wild type accessions and expressivity of *amiPPD*. (A) Venn diagram showing the number of genes having their basal expression levels correlating with 'rosette' and/or 'selective leaf' expressivity of *amiPPD*. Numbers in bracket correspond to the number of genes having their expression levels showing either a positive correlation (first number) or negative correlation (second number). (B) Venn diagram representing the number of genes overlapping between genes having their basal expression levels correlating with 'rosette' and/or 'selective leaf' expressivity of *amiPPD* and DE genes for *amiPPD.* (C) Over-representing MapMan categories for genes having their expression correlating with 'rosette' and/or 'selective leaf' expressivity of *amiPPD*. (D) Expression level of genes correlating with 'rosette' expressivity of *amiPPD.* Values are count per million from RNA sequencing data. A heat map of 'rosette' expressivity of 13 wild type accessions is shown at the top.

Discussion

It has been shown that different individuals having the same mutation can show different phenotype severity (Dowell et al., 2010; Chandler et al., 2014; Vu et al., 2015). Many factors are involved in this phenomenon and one of them is a genetic component. In this study, we examined how different genetic backgrounds could influence the effect of different genetic perturbations in Arabidopsis natural accessions. We used three different transgenes, overexpression of *GA20ox1*, overexpression of an amiRNA targeting *PPD2* and overexpression of a dominant negative form of *DA1,* that were previously shown to promote leaf growth in Col-0 background. These three transgenes were introduced in 17 different Arabidopsis accessions including Col-0 and the effect on leaf growth of each transgene was measured. A comparative analysis of the effect of each transgene in the different genetic backgrounds showed that the expressivity observed is transgene specific and also background specific. However, the phenotype differences seem to not be due to sequence diversity between the transgenes and the endogenous gene in the different accessions. We also found that the effects of *GA20ox1*^{OE} in the 17 Arabidopsis accessions were more diverse compared to the effects of *amiPPD* and *DN-DA1*^{OE} transgenic lines. Although in most cases, younger leaves are positively affected by *GA20ox1*^{OE}, due to a total number of leaves varying in the different accessions, the leaf position of "younger leaves" was somewhat subjective depending on the accession. For instance, in the accession Sha the younger leaves which are leaf 5, 6, 7 and 8 were positively affected in the *GA20ox1* ^{OE} transgenic lines whereas the younger leaves of transgenics in Yeg-1 corresponded to leaf 9, 10, 11, 12, 13 and 14. Therefore, 'selective leaf' expressivity of *GA20ox1*^{OE} was much more variable than 'selective leaf' expressivity of the two other transgenes. Furthermore, some transgenic lines showed almost the entire rosette leaves affected by *GA20ox1*^{OE} either positively or negatively while transgenic lines of *amiPPD* and *DN-DA1*^{OE} showed relatively consistent positive effect on older leaves and negative effect on younger leaves. A possible reason to explain this difference is that *GA20ox1*^{OE} causes the formation of high levels of gibberellins that are likely to affect more broad developmental processes compared to the other transgenes. PPD2 is a negative regulator of transcription (White, 2006; Gonzalez et al., 2015) and as such closer to executor genes as compared to GA20ox1. The molecular function of DA1 is less clear but genetic evidence showed its involvement in proteolysis of the transcription factor TCP14/15 (Peng et al., 2015) thus again more downstream in a signalling cascade as compared to GA20ox1. While *amiPPD* and *DN-DA1*^{OE} affect mainly pavement cell number, transgenic lines of *GA20ox1*^{OE} showed alteration in pavement cell number and/or area. Interestingly, the correlation analysis between wild type phenotypes and the expressivity of the transgenes showed different results depending on the transgenes. The expressivity values of *amiPPD* and *DN-DA1*^{OE} were negatively correlated with growth related parameters of wild type (biomass, rosette area and/or total leaf number). However, these correlations were positive for *GA20ox1*^{OE}. These findings demonstrate that there is no general rule explaining how wild type plant is affected by a transgene, but the outcome seems to be dependent on which transgene is introduced in which accession.

Most of the genes expressed in wild-type accessions that were correlating with the expressivity of the two transgenes (*GA20ox1*^{OE} and *amiPPD*) were not differentially expressed in the transgenic lines but belonged to similar functional categories. For instance, we found genes that are involved in hormone metabolism negatively correlating with 'rosette' expressivity of *GA20ox1*^{OE}. One of the member of the SAUR-like auxin responsive protein family was found and other members of this family were found in the DE gene set of *GA20ox1*^{OE}. Besides, a gene involved in ethylene-activated signalling pathway, *MBF1C* (*MULTIPROTEIN BRIDGING FACTOR1C*) encoding a transcription co-activator was also found to be negatively correlated. It is known that the activation of ethylene signaling suppresses accumulation of bio-active GA forms (Achard et al., 2007). Ethylene act as external factors that negatively regulate GA biosynthesis (Sun, 2011). If an accession has low expression of ethylene related genes, the expressivity could be high due to less antagonistic action of ethylene towards GA. This result suggests a tight interaction between GA and other hormones (Weiss and Ori, 2007) which could influence leaf growth upon overexpression of *GA20ox1*. The expression level of *FLC* encoding a transcription factor negatively affecting flowering time (Deng et al., 2011) was positively correlated with expressivity of *GA20ox1*^{OE}. *FLC* was highly expressed in wild types showing relatively late flowering phenotype while other accessions which have an early flowering phenotype showed low expression of this gene. It has been shown that FLC delays the progression from juvenile-to-adult phase (Willmann and Poethig, 2011). A recent study has shown that DELLAs interact with FLC to enhance the transcriptional inhibition ability of FLC to its target genes, *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1)* and *FLOWERING LOCUS T (FT)* to regulate flowering time (Li et al., 2015). Thus, it could be that due to the high basal level of *FLC* in some accessions, the effect on flowering time of high gibberellin levels in the transgenics is counteracted leading to the production of same amount of leaves as in wild type that can grow more. On the other hand, in accessions with low levels of *FLC* transcripts, the overexpression of *GA20ox1* leads to early flowering resulting in the production of less leaves that grow less.

We could also identify in wild type, genes having expression levels correlating with expressivity of *amiPPD*. Some of these genes are direct targets of PPD2. Also protein sequence diversity was found in the PPD2 protein of four accessions (Blh-1, ICE138, ICE153, and ICE163) compared to Col-0. This sequence divergence might lead to the production of PPD2 proteins having different affinity to promoter regions of target genes or different ability to interact with KIX8/9 proteins and the corepressor TOPLESS (Gonzalez et al., 2015) in function of the accession. Therefore it is possible that in different genetic backgrounds, the different variants of *PPD2* regulate their gene targets differently therefore causing distinct effects on the regulated networks upon perturbation.

There have been several studies on the effect of mutations modulated by genetic backgrounds. In *C. elegans*, 20% of the genes showed a genetic background-specific phenotype in a loss-of-function screening analysis in two different genetic backgrounds (Vu et al., 2015). In our study, we could observe that three transgenes in 17 Arabidopsis accessions triggered variable phenotypic outcomes on leaf growth. In summary, the effect of transgenes is depending on the

combination between transgene and genetic background probably due to epistatic interactions with other genes in the same or different pathways. We found that natural variation of basal expression of some genes in wild type possibly could influence the effect of a transgene highlighting these potential epistatic effects. The putative role of these genes in affecting the expressivity of the transgenes studied could be discovered by altering their expression in the transgenic lines.

Methods

Plant material and growth conditions

Seventeen Arabidopsis accessions were selected to cover a large genetic variation of *Arabidopsis thaliana* (Table 4.1) and used to generate lines overexpressing a dominant negative form of *DA1*. A full cDNA sequence from Col-0 harboring the $DA1^{R358K}$ (Li et al., 2008) was cloned under the control of the 35S CAMV promoter in the fluorescence-accumulating seed technology (FAST) vectors (Shimada et al., 2010) and introduced into the 17 accessions following the floral dip protocol (Clough and Bent, 1998). Dried transgenic T_1 seeds were selected based on fluorescence signal in the seed coat and sown on soil for seed production. T_2 transgenic seeds were harvested and selection of five independent single-locus insertion lines (75% of fluorescent seeds) was done. Seeds were sown on soil for seed production and expression of the transgene was verified by RT-qPCR. From these lines, at least two and maximum five independent T_3 homozygote lines for each accession were selected for further experiments. All plants were grown MIRGIS platform in soil under a 16-hours-day/8-hours-night regime at 21°C in a growth chamber. The daily images of plants obtained from MIRGIS are analyzed in Chapter 8.

Phenotypic analysis

Leaf series data were obtained from 25-day-old plants by dissecting individual leaves (from cotyledons to the younger rosette leaves) and mounting them on a 1 % agar plate, and the leaf area was measured with the ImageJ software (http://rsb.info.nih.gov/ij/). The leaf series for the transgenic lines were performed in one repeat by growing the transgenic lines together with their corresponding wild type.

For the leaf series data, leaf area was log transformed to stabilize the variance. The mean model consisted of the main effects of *DN-DA1*^{OE} on leaf size and their interaction term. Due to the unbalanced and complex nature of the data, the Kenward-Rogers approximation for computing the denominator degrees of freedom for the tests of fixed effects was used. An autoregressive structure was used to model the correlations between measurements done on the leaves originating from the same plant. The main interest was in the effect of the gene on leaf area for each leaf separately. Simple tests of effects were performed at each leaf between the transgenic lines and the corresponding wild type. Difference estimates were represented as % to

the least-square means estimate of the wild type and leaf. Separate models were run for each accession as they were grown in separate experiments. For each experiment, the data was truncated so that there were at least two observations for each leaf of both the transgenic lines and the corresponding wild type. The analysis was performed with the mixed and plm procedure of SAS [Version 9.4 of the SAS System for windows 7 64bit. Copyright © 2002-2012 SAS Institute Inc. Cary, NC, USA (www.sas.com). Residual diagnostics were carefully examined.

The expressivity of 'selective leaf' was determined by averaging the ratio of leaf 1/2, 3 and 4 for each leaf. To calculate the 'selective leaf' expressivity per accession, the 'selective leaf' expressivity per transgenic line were taken for the average estimated leaf effect. 'Rosette' expressivity was estimated as a ratio of the wild-type rosette area to that of a transgenic line. In case of 'rosette' expressivity per accession, the mean of 'rosette' expressivity per transgenic line of an accession has been taken.

Cellular analysis was done as previously described (Andriankaja et al., 2012) in three repeats. For transgenic lines, significant differences were determined with a two-way ANOVA test with genotype and repeat as main factors in R software (v 3.0.1)(R Core Team, 2015). Differences between the wild type and corresponding transgenic lines were estimated and declared significant when the adj-P value < 0.05 with Tukey's method.

Sequence extraction and alignment

The different read libraries were quality checked using FastQC v0.9.1 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapters were trimmed using cutadapt in python v3.1.1 with the options: --adapter 'GATCGGAAGAGCACACGTCTGAACTCCAGTCAC', --overlap 10 and --minimum-length 35 (Martin, 2011). Quality filtering was performed using Fastx v0.0.13 with the options: -Q 33, -q 10 and -p 75 (http://hannonlab.cshl.edu/fastx_toolkit/). Reads were also trimmed using Fastx v0.0.13 with the options: -Q 33, -t 20 and -l 35. Forward and reverse reads were subsequently collapsed into a single file. After preprocessing, the different read libraries were mapped to the TAIR10 genome using gsnap v2013-02-05 with the options --trim-mismatch-score -3, -k 15, -A sam, -B 4, -n 50, -w 15000, -a off, --quality-protocol sanger, --pairmax-rna 15000, -N 1 and -m 5 (Wu and Nacu, 2010). Only uniquely mapping reads were further considered. Next, sorting and deduplication of the read libraries was performed using picard v1.129 (http://broadinstitute.github.io/picard/). GATK v3.3.0 was used for variant calling (Van der Auwera et al., 2013). Analysis was based on recommendations in 'Best practices for RNA-seq' (https://www.broadinstitute.org/gatk/guide/best-practices?bpm=RNAseq). Before variant calling was performed, the different libraries were preprocessed using the tools splitnCigar, haplotypecaller, realignertargetcreator, indelrealigner, baserecalibrator and printreads. In the haplotypecaller step only high quality scores were considered by setting a quality of 50. Next, a

multi-sample variant calling was performed using haplotypecaller. In this step, all samples are analysed together. Variants were filtered using VariantFiltration with the options -window 35, -cluster 3, -filterName FS, -filter "FS > 30.0", -filterName QD and -filter "QD < 2.0". The resulting variants file was split by sample using bcftools (http://github.com/samtools/bcftools). Sequences were extracted for the genes (AT4G14713, AT4G14720, AT1G19270, AT2G01570 and AT4G25420) using the alternative alleles for each sample using the GATK tool Fasta Alternate Reference Maker (Van der Auwera et al., 2013) and based on the CDS coordinates (based on the structural annotation of TAIR10). The reverse complement was generated for genes located at the negative strand and subsequently protein sequences were extracted using custom scripts.

To align extracted sequence, CLC main Workbench 6.0 was used (CLC bio, a QIAGEN Company; http://www.clcbio.com/).

Visualization of the RNA sequencing data

Enrichment analysis was done using mapman categories (Ramšak et al., 2014) (http://bar.utoronto.ca/ntools/cgi-bin/ntools_classification_superviewer.cgi) with DE genes filtered for an FDR value lower than 0.05.

Heat maps are generated in Mev (v 4.9)(Howe et al., 2011) for DE gene filtered for an FDR value lower than 0.05 and a 1.5-fold change threshold between transgenic lines and the wild type. Hierarchical clustering was done for both genes and samples with Manhattan distance metrics in Mev (v 4.9)(Howe et al., 2011).

Generating phylogeny tree

The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The optimal tree with the sum of branch length = 2.19379816 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Kimura 2-parameter method (Kimura, 1980) and are in the units of the number of base substitutions per site. The analysis involved 14 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 130117 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al., 2011).

Correlation analysis

Pearson correlation coefficients were calculated with corr.test function in R. The adjusted P values of correlations were calculated with a permutation test. For each analysis, namely, the phenotype-phenotype and expressivity-RNA-seq count data (filtered out genes expressed less than either 10 (for RNA seq data from *GA20ox1*^{OE} lines) or 30 (for RNA seq data from *amiPPD* lines) counts in every samples) we ran 1000 permutations. Each permutation run

involves a permutation of a single phenotype or expressivity, respectively to analysis, and calculation of correlations between permuted vector and the rest of analysed data. The adjusted P values over all permutation runs per analysis were calculated as a proportion of correlation coefficients correlated in a higher degree than a tested correlation (r) to the number of permuted correlations (n); with a formula (r+1)/(n+1) (North et al., 2002). The significant correlations, FDR<0.05, were visualized in Cytoscape (Cline et al., 2007).

References

Achard, P., Baghour, M., Chapple, A., Hedden, P., Van Der Straeten, D., Genschik, P., Moritz, T., and Harberd, N.P. (2007). The plant stress hormone ethylene controls floral transition via DELLA-dependent regulation of floral meristem-identity genes. Proc Natl Acad Sci U S A **104:** 6484-6489.

Andriankaja, M., Dhondt, S., De Bodt, S., Vanhaeren, H., Coppens, F., De Milde, L., Mühlenbock, P., Skirycz, A., Gonzalez, N., Beemster, G.T.S., and Inzé, D. (2012). Exit from proliferation during leaf development in *Arabidopsis thaliana*: a not-so-gradual process. Dev. Cell **22:** 64-78.

Baylis, T., Cierlik, I., Sundberg, E., and Mattsson, J. (2013). SHORT INTERNODES/STYLISH genes, regulators of auxin biosynthesis, are involved in leaf vein development in Arabidopsis thaliana. The New phytologist **197:** 737-750.

Chandler, C.H., Chari, S., and Dworkin, I. (2013). Does your gene need a background check? How genetic background impacts the analysis of mutations, genes, and evolution. Trends Genet. **29:** 358-366.

Chandler, C.H., Chari, S., Tack, D., and Dworkin, I. (2014). Causes and consequences of genetic background effects illuminated by integrative genomic analysis. Genetics **196:** 1321-1336.

Cline, M.S., et al. (2007). Integration of biological networks and gene expression data using Cytoscape. Nat. Protoc. **2:** 2366-2382.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. **16:** 735-743.

Deng, W., Ying, H., Helliwell, C.A., Taylor, J.M., Peacock, W.J., and Dennis, E.S. (2011). FLOWERING LOCUS C (FLC) regulates development pathways throughout the life cycle of Arabidopsis. Proc Natl Acad Sci U S A **108:** 6680-6685.

Dowell, R.D., et al. (2010). Genotype to phenotype: a complex problem. Science **328:** 469.

Gonzalez, N., et al. (2010). Increased leaf size: different means to an end. Plant Physiol. **153:** 1261-1279.

Gonzalez, N., et al. (2015). A Repressor Protein Complex Regulates Leaf Growth in Arabidopsis. Plant Cell **27:** 2273-2287.

Howe, E.A., Sinha, R., Schlauch, D., and Quackenbush, J. (2011). RNA-Seq analysis in MeV. Bioinformatics **27:** 3209-3210.

Hsu, P.K., and Tsay, Y.F. (2013). Two phloem nitrate transporters, NRT1.11 and NRT1.12, are important for redistributing xylem-borne nitrate to enhance plant growth. Plant Physiol **163:** 844-856.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16:** 111-120.

Li, M., An, F., Li, W., Ma, M., Feng, Y., Zhang, X., and Guo, H. (2015). DELLA Proteins Interact with FLC to Repress the Flowering Transition. J Integr Plant Biol.

Li, Y., Zheng, L., Corke, F., Smith, C., and Bevan, M.W. (2008). Control of final seed and organ size by the *DA1* gene family in *Arabidopsis thaliana*. Genes Dev. **22:** 1331-1336.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal **17:** 10-12.

North, B.V., Curtis, D., and Sham, P.C. (2002). A note on the calculation of empirical *P* values from Monte Carlo procedures. Am. J. Hum. Genet. **71:** 439-441.

Paaby, A.B., and Rockman, M.V. (2014). Cryptic genetic variation: evolution's hidden substrate. Nat Rev Genet **15:** 247-258.

Peng, Y., Chen, L., Lu, Y., Wu, Y., Dumenil, J., Zhu, Z., Bevan, M.W., and Li, Y. (2015). The ubiquitin receptors DA1, DAR1, and DAR2 redundantly regulate endoreduplication by modulating the stability of TCP14/15 in Arabidopsis. Plant Cell **27:** 649-662.

R Core Team. (2015). R: a language and environment for statistical computing. Vienna, Austria (http://www.R-project.org/).

Ramšak, Ž., Baebler, Š., Rotter, A., Korbar, M., Mozetič, I., Usadel, B., and Gruden, K. (2014). GoMapMan: integration, consolidation and visualization of plant gene annotations within the MapMan ontology. Nucleic Acids Res. **42:** D1167-1175.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol **4:** 406-425.

Shimada, T.L., Shimada, T., and Hara-Nishimura, I. (2010). A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. Plant J. **61:** 519-528.

Sun, T.-p. (2011). The molecular mechanism and evolution of the GA-GID1-DELLA signaling module in plants. Curr. Biol. **21:** R338-R345.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28:** 2731-2739.

Tominaga, R., Iwata, M., Sano, R., Inoue, K., Okada, K., and Wada, T. (2008). Arabidopsis CAPRICE-LIKE MYB 3 (CPL3) controls endoreduplication and flowering development in addition to trichome and root hair formation. Development **135:** 1335-1345.

Van der Auwera, G.A., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics **11:** 11 10 11-11 10 33.

Vu, V., Verster, A.J., Schertzberg, M., Chuluunbaatar, T., Spensley, M., Pajkic, D., Hart, G.T., Moffat, J., and Fraser, A.G. (2015). Natural variation in gene expression modulates the severity of mutant phenotypes. Cell **162:** 391-402.

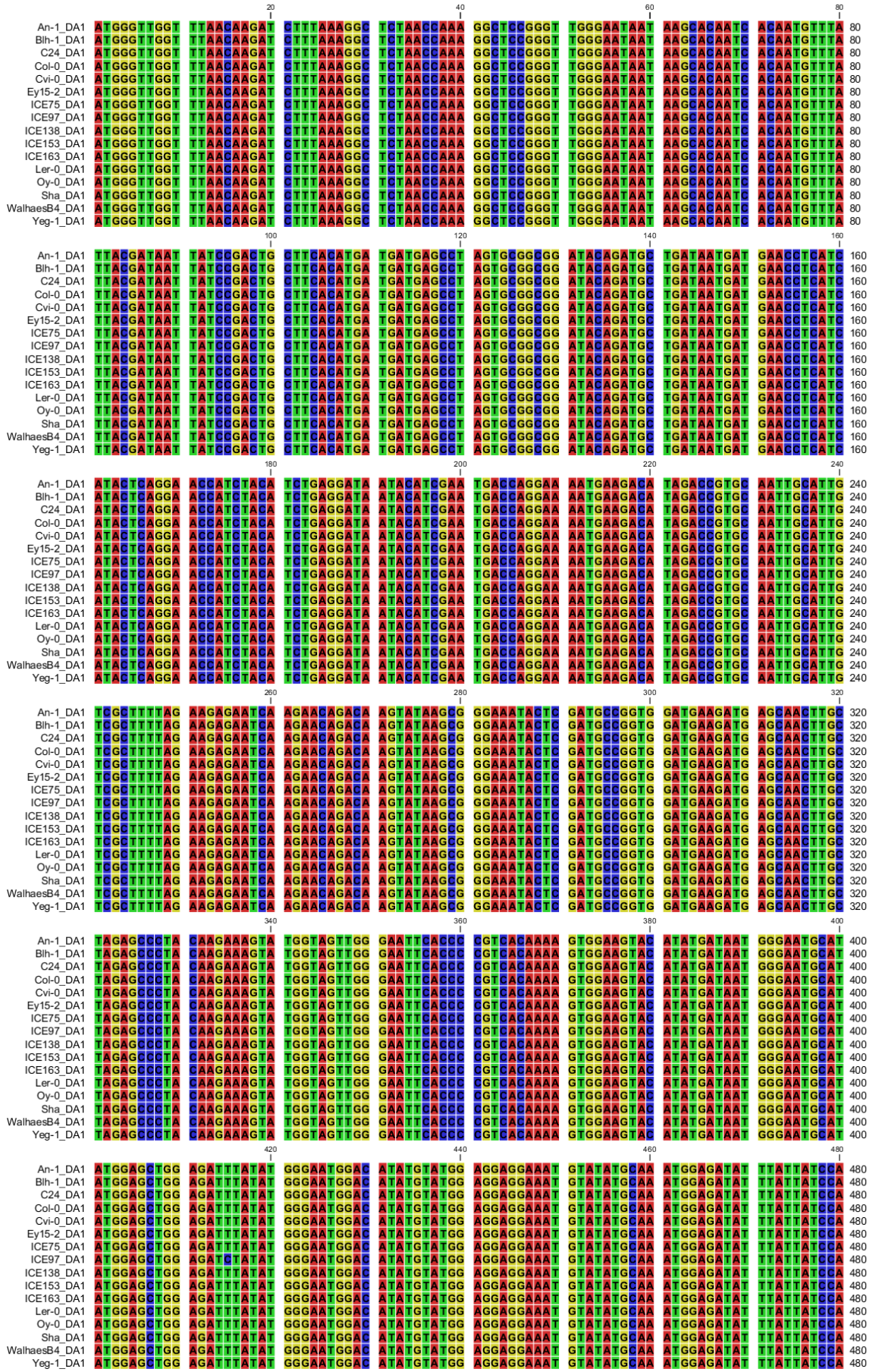
Weiss, D., and Ori, N. (2007). Mechanisms of cross talk between gibberellin and other hormones. Plant Physiol. **144:** 1240-1246.

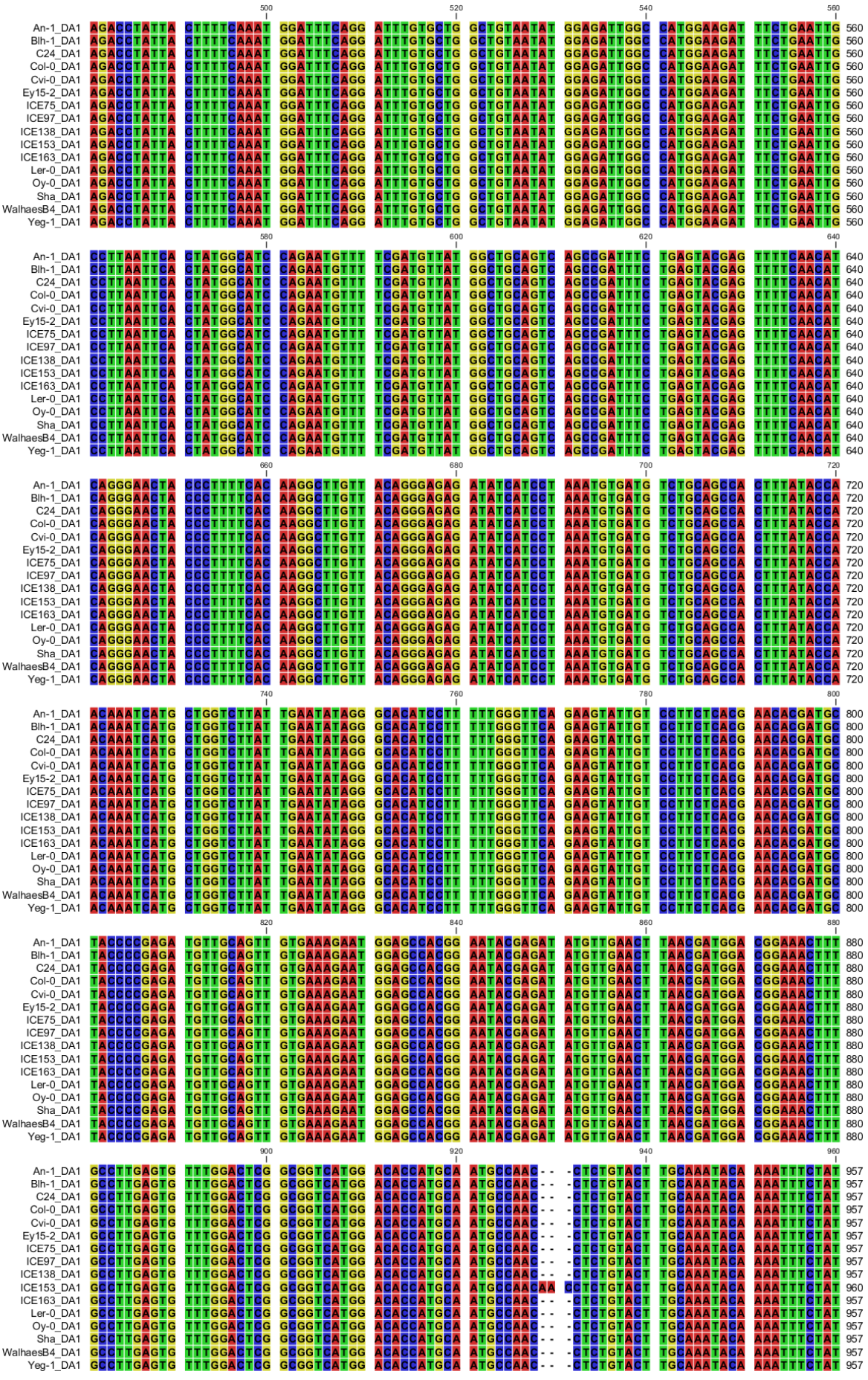
White, D.W.R. (2006). *PEAPOD* regulates lamina size and curvature in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **103:** 13238-13243.

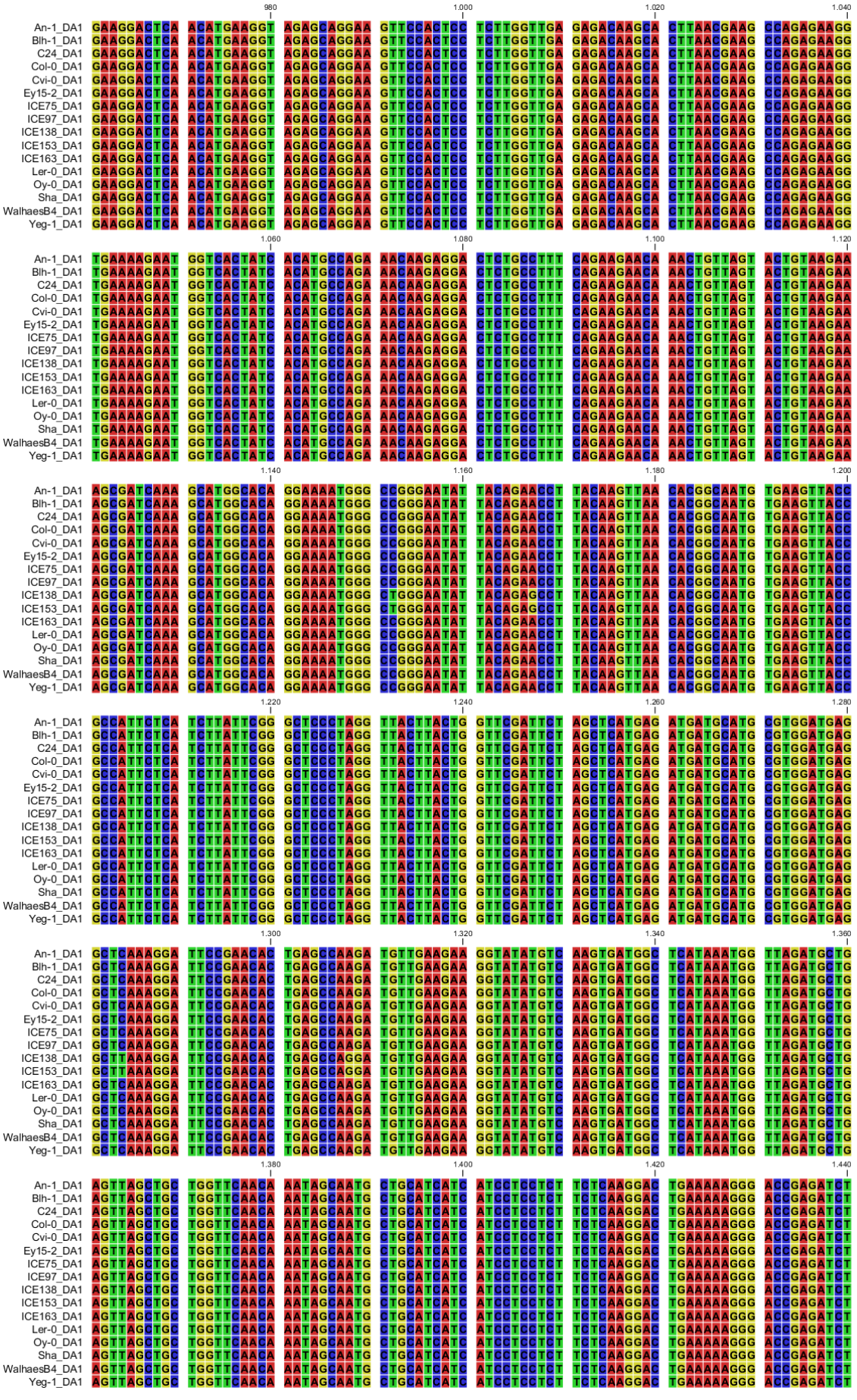
Willmann, M.R., and Poethig, R.S. (2011). The effect of the floral repressor FLC on the timing and progression of vegetative phase change in Arabidopsis. Development **138:** 677-685.

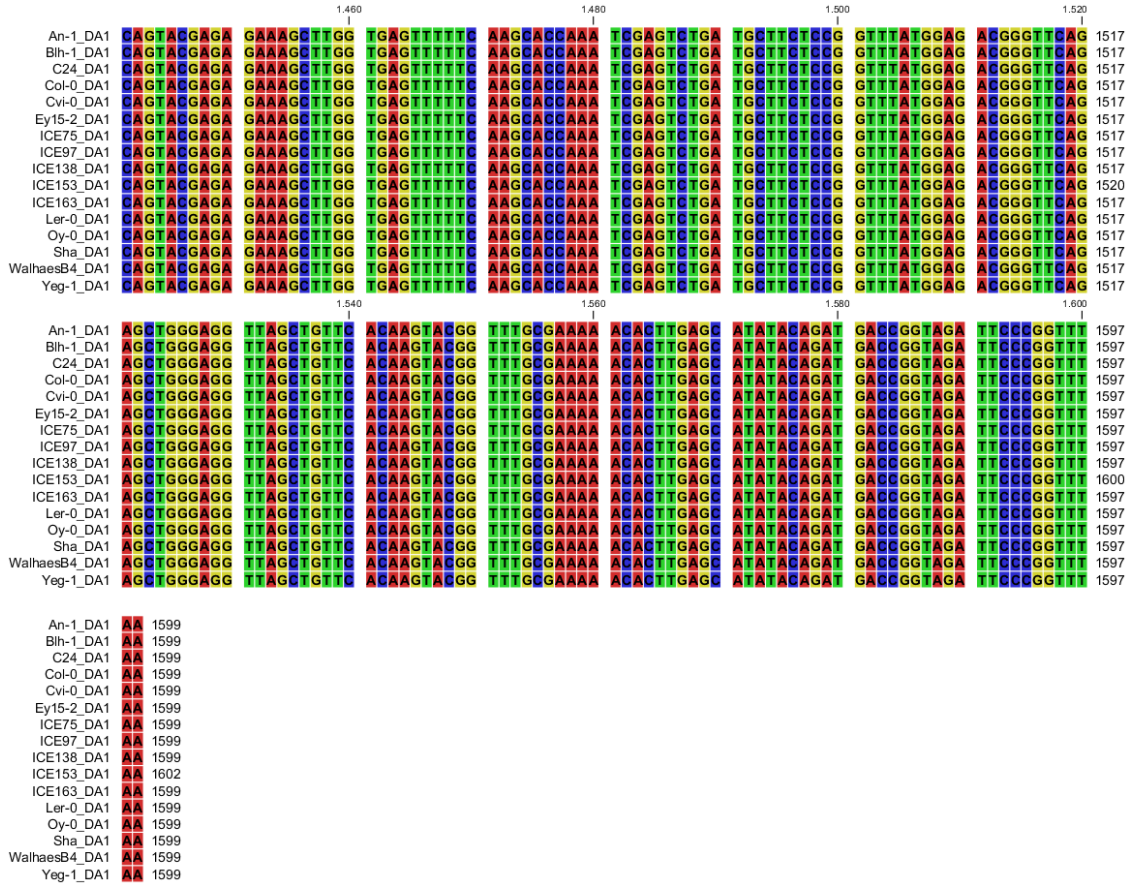
Wu, T.D., and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics **26:** 873-881.

A

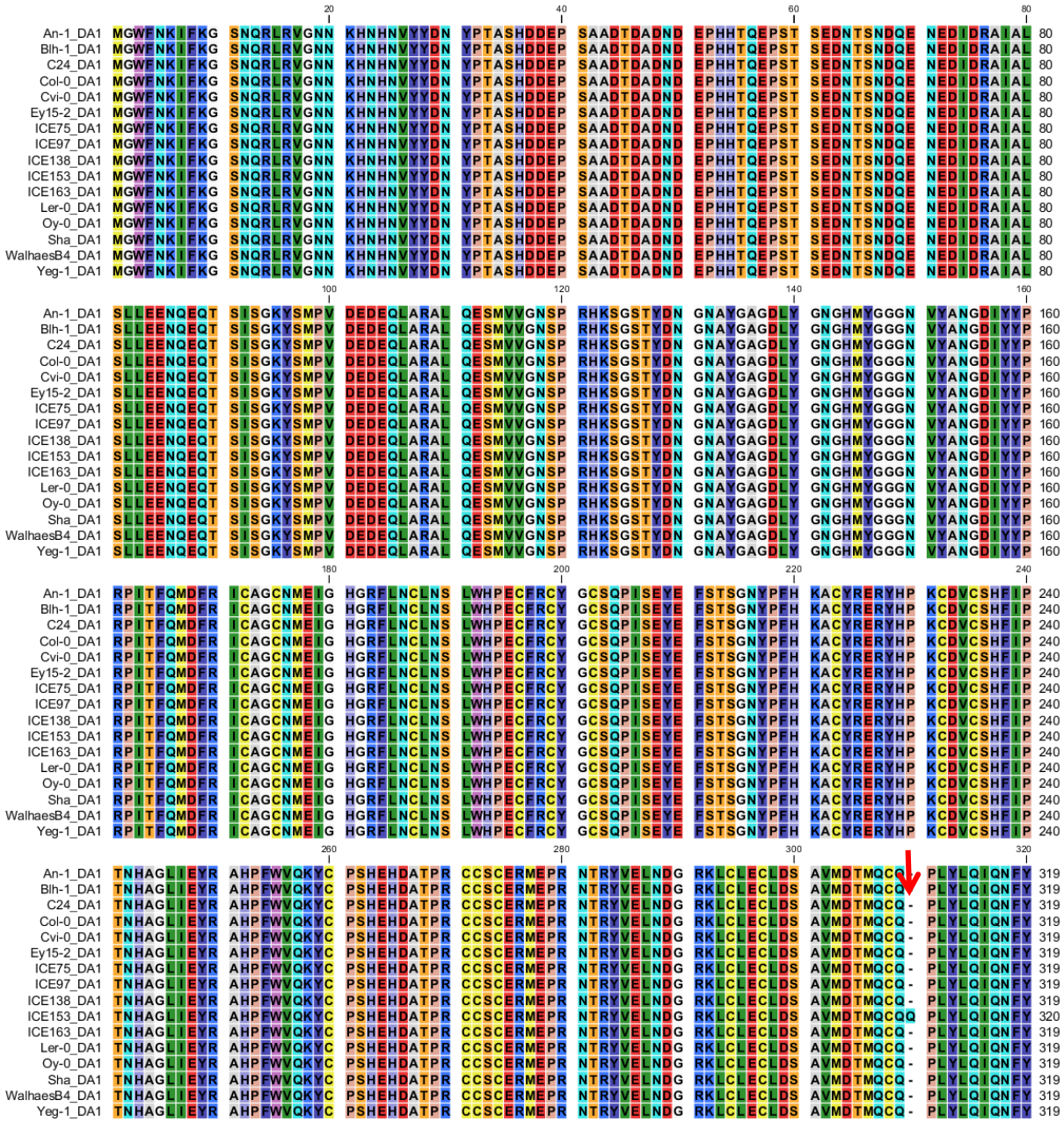

An-1_DA1
Blh-1_DA1
C24_DA1
Col-0_DA1
Cvi-0_DA1
Ey15-2_DA1
ICE75_DA1
ICE97_DA1
ICE138_DA1
ICE153_DA1
ICE163_DA1
Ler-0_DA1
Oy-0_DA1
Sha_DA1
WalhaesB4_DA1
Yeg-1_DA1

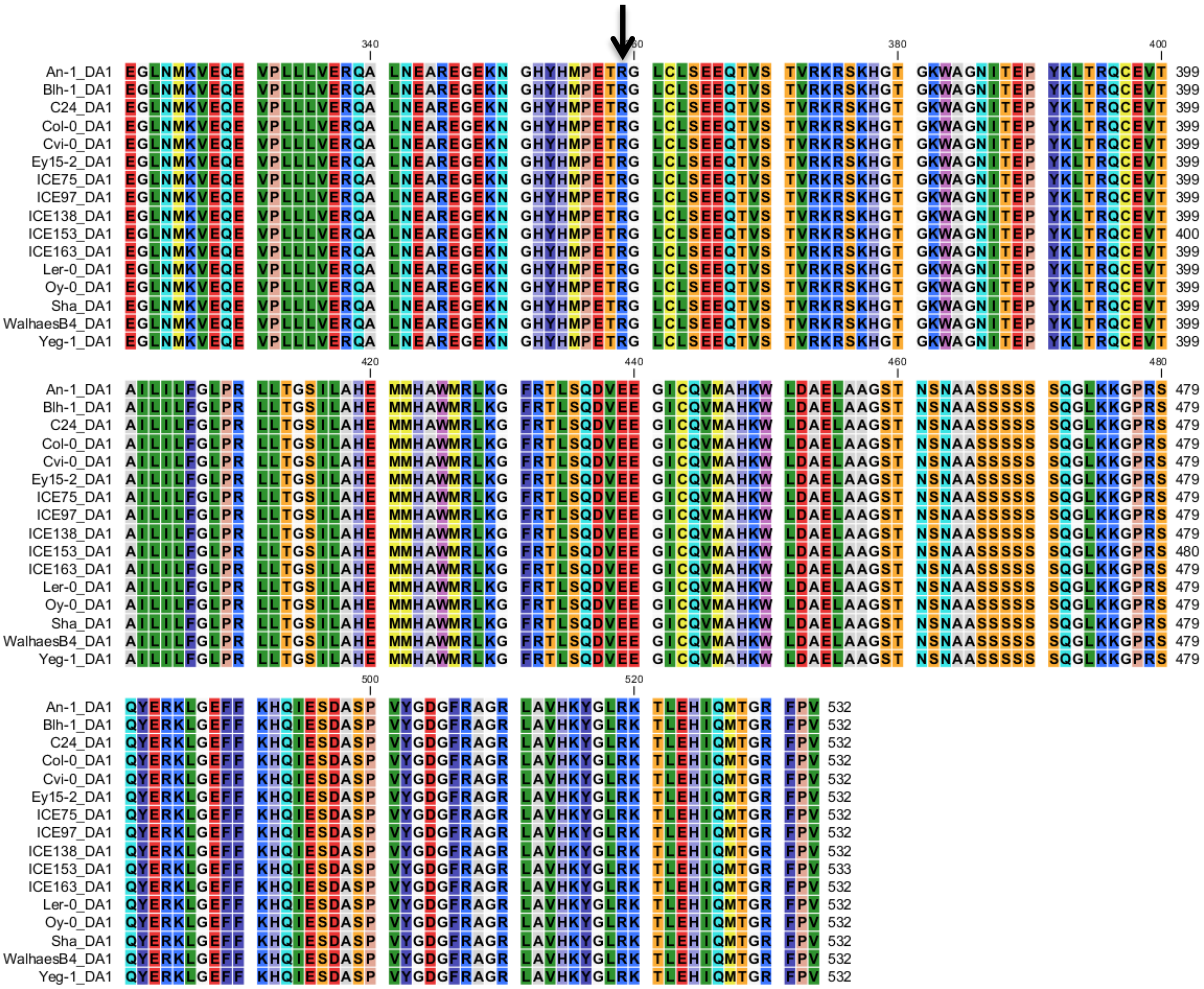

An-1_DA1
Blh-1_DA1
C24_DA1
Col-0_DA1
Cvi-0_DA1
Ey15-2_DA1
ICE75_DA1
ICE97_DA1
ICE138_DA1
ICE153_DA1
ICE163_DA1
Ler-0_DA1
Oy-0_DA1
Sha_DA1
WalhaesB4_DA1
Yeg-1_DA1


980 1.000 1.020 1.040
An-1_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
Blh-1_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
C24_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
Col-0_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
Cvi-0_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
Ey15-2_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
ICE75_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
ICE97_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
ICE138_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
ICE153_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1040
ICE163_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
Ler-0_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
Oy-0_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
Sha_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
WalhaesB4_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
Yeg-1_DA1 GAAGGACTCA ACATGAAGGT AGAGCAGGAA GTTCCACTCC TCTTGGTTGA GAGACAAGCA CTTAACGAAG CCAGAGAAGG 1037
1.060 1.080 1.100 1.120
An-1_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
Blh-1_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
C24_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
Col-0_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
Cvi-0_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
Ey15-2_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
ICE75_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
ICE97_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
ICE138_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
ICE153_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1120
ICE163_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
Ler-0_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
Oy-0_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
Sha_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
WalhaesB4_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
Yeg-1_DA1 TGAAAAGAAT GGTCACTATC ACATGCCAGA AACAAGAGGA CTCTGCCTTT CAGAAGAACA AACTGTTAGT ACTGTAAGAA 1117
1.140 1.160 1.180 1.200
An-1_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
Blh-1_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
C24_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
Col-0_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
Cvi-0_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
Ey15-2_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
ICE75_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
ICE97_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
ICE138_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CTGGGAATAT TACAGAGCCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
ICE153_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CTGGGAATAT TACAGAGCCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1200
ICE163_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CGGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
Ler-0_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
Oy-0_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
Sha_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
WalhaesB4_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
Yeg-1_DA1 AGGGATCAAA GCATGGCACA GGAAAATGGG CCGGGAATAT TACAGAACCT TACAAGTTAA CACGGCAATG TGAAGTTACC 1197
1.220 1.240 1.260 1.280
An-1_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
Blh-1_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
C24_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
Col-0_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
Cvi-0_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
Ey15-2_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
ICE75_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
ICE97_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
ICE138_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
ICE153_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1280
ICE163_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
Ler-0_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
Oy-0_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
Sha_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
WalhaesB4_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
Yeg-1_DA1 GCCATTCTCA TCTTATTCGG GCTCCCTAGG TTACTTACTG GTTCGATTCT AGCTCATGAG ATGATGCATG CGTGGATGAG 1277
1.300 1.320 1.340 1.360
An-1_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
Blh-1_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
C24_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
Col-0_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
Cvi-0_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
Ey15-2_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
ICE75_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
ICE97_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
ICE138_DA1 GCTTAAAGGA TTCCGAACAC TGAGCCAGGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
ICE153_DA1 GCTTAAAGGA TTCCGAACAC TGAGCCAGGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1360
ICE163_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
Ler-0_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
Oy-0_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
Sha_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
WalhaesB4_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
Yeg-1_DA1 GCTCAAAGGA TTCCGAACAC TGAGCCAAGA TGTTGAAGAA GGTATATGTC AAGTGATGGC TCATAAATGG TTAGATGCTG 1357
1.380 1.400 1.420 1.440
An-1_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
Blh-1_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
C24_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
Col-0_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
Cvi-0_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
Ey15-2_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
ICE75_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
ICE97_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
ICE138_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
ICE153_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1440
ICE163_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
Ler-0_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
Oy-0_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
Sha_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
WalhaesB4_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437
Yeg-1_DA1 AGTTAGCTGC TGGTTCAACA AATAGCAATG CTGCATCATC ATCCTCCTCT TCTCAAGGAC TGAAAAAGGG ACCGAGATCT 1437


1.460
1.480
1.500
1.520
An-1_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCCG GTTTATGGAG ACGGGTTCAG 1517
Blh-1_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCGG GTTTATGGAG ACGGGTTCAG 1517
C24_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCGG GTTTATGGAG ACGGGTTCAG 1517
Col-0_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCGG GTTTATGGAG ACGGGTTCAG 1517
Cvi-0_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCCG GTTTATGGAG ACGGGTTCAG 1517
Ey15-2_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCCG GTTTATGGAG ACGGGTTCAG 1517
ICE75_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCGG GTTTATGGAG ACGGGTTCAG 1517
ICE97_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCGG GTTTATGGAG ACGGGTTCAG 1517
ICE138_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCGG GTTTATGGAG ACGGGTTCAG 1517
ICE153_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCCG GTTTATGGAG ACGGGTTCAG 1520
ICE163_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCCG GTTTATGGAG ACGGGTTCAG 1517
Ler-0_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCGG GTTTATGGAG ACGGGTTCAG 1517
Oy-0_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCGG GTTTATGGAG ACGGGTTCAG 1517
Sha_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCGG GTTTATGGAG ACGGGTTCAG 1517
WalhaesB4_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCGG GTTTATGGAG ACGGGTTCAG 1517
Yeg-1_DA1 CAGTACGAGA GAAAGCTTGG TGAGTTTTTC AAGCACCAAA TCGAGTCTGA TGCTTCTCGG GTTTATGGAG ACGGGTTCAG 1517
1.540
1.560
1.580
1.600
An-1_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
Blh-1_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
C24_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
Col-0_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
Cvi-0_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
Ey15-2_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
ICE75_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
ICE97_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
ICE138_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
ICE153_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1600
ICE163_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
Ler-0_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
Oy-0_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
Sha_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
WalhaesB4_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
Yeg-1_DA1 AGCTGGGAGG TTAGCTGTTC ACAAGTACGG TTTGCGAAAA ACACTTGAGC ATATACAGAT GACCGGTAGA TTCCCGGTTT 1597
An-1_DA1 AA 1599
Blh-1_DA1 AA 1599
C24_DA1 AA 1599
Col-0_DA1 AA 1599
Cvi-0_DA1 AA 1599
Ey15-2_DA1 AA 1599
ICE75_DA1 AA 1599
ICE97_DA1 AA 1599
ICE138_DA1 AA 1599
ICE153_DA1 AA 1602
ICE163_DA1 AA 1599
Ler-0_DA1 AA 1599
Oy-0_DA1 AA 1599
Sha_DA1 AA 1599
WalhaesB4_DA1 AA 1599
Yeg-1_DA1 AA 1599

B


An-1_DA1
Blh-1_DA1
C24_DA1
Col-0_DA1
Cvi-0_DA1
Ey15-2_DA1
ICE75_DA1
ICE97_DA1
ICE138_DA1
ICE153_DA1
ICE163_DA1
Ler-0_DA1
Oy-0_DA1
Sha_DA1
WalhaesB4_DA1
Yeg-1_DA1



C

	GA20ox1		PPD		DA1	
Accession	cDNA	Protein	cDNA	Protein	cDNA	Protein
An-1	0	0	0.11	0	0	0
Blh-1	0.09	0	0.21	0.32	0	0
C24	0.09	0	0.11	0	0	0
Cvi-0	0	0	0.11	0	0	0
Ey15-2	0.17	0	0.11	0	0	0
ICE75	0.17	0	\	\	0	0
ICE97	0.09	0	\	\	0.06	0
ICE138	0.09	0	0.21	0.32	0.25	0
ICE153	0	0	0.21	0.32	0.44	0.18
ICE163	0.09	0	0.32	0.63	0	0
Ler-0	0.17	0	0	0	0	0
Oy-0	0.09	0	0	0	0	0
Sha	0.17	0	0.11	0	0	0
WalhaesB4	0	0	0	0	0	0
Yeg-1	0	0	0	0	0	0

Figure 7.4 Sequence alignments of the three endogenous genes and their corresponding proteins used in this project in 17 *Arabidopsis* accessions. (A and B) cDNA and protein sequence of DA1. The red arrow points the location of sequence divergence in ICE153 and the blue arrow indicates where the point mutation in *DN-DA1*^{OE} transgene is located. (C) Percentage differences of sequences of the three genes (GA20ox1, PPD2, and DA1) and proteins compared to the sequences of Col-0. \, no transgenic plant.

Table 7.1 List of genes having their expression correlating with rosette penetrance of *GA20ox1*^{OE} in wild type. Only genes involved in transport, hormone metabolism, and RNA MapMan categories are shown. The expression levels of the underlined genes is shown in Figure 7.6D.

AT code	Correlation	adjP-value		ManMan category	Gene description
At1g52190	-0.81	0.001	Transport	peptides and oligopeptides	AtNPF1.2_NPF1.2_NRT1.11__Major facilitator superfamily protein
At1g16390	-0.74	0.004	Transport	unspecified cations	ATOCT3_OCT3__organic cation/carnitine transporter 3
At3g62270	0.74	0.016	Transport	unspecified anions	BOR2__HCO3- transporter family
At5g14870	0.77	0.011	Transport	cyclic nucleotide or calcium regulated channels	ATCNGC18_CNGC18__cyclic nucleotide-gated channel 18
At2g03260	-0.70	0.006	Transport	phosphate	EXS (ERD1/XPR1/SYG1) family protein
At4g28390	-0.75	0.003	Transport	metabolite transporters at the mitochondrial membrane	AAC3_ATAAC3__ADP/ATP carrier 3
At1g44350	-0.71	0.005	Hormone metabolism	auxin.synthesis-degradation	ILL6__IAA-leucine resistant (ILR)-like gene 6
At1g29500	-0.72	0.005	Hormone metabolism	auxin.induced-regulated-responsive-activated	SAUR66__SAUR-like auxin-responsive protein family
At3g24500	-0.71	0.005	Hormone metabolism	ethylene.induced-regulated-responsive-activated	ATMBF1C_MBF1C__multiprotein bridging factor 1C
At4g36470	-0.83	0.001	Hormone metabolism	salicylic acid.synthesis-degradation	S-adenosyl-L-methionine-dependent methyltransferases superfamily protein
At1g14640	-0.72	0.004	RNA	processing	SWAP (Suppressor-of-White-APricot)/surp domain-containing protein
At3g54490	-0.71	0.005	RNA	transcription	RPB5E__RNA polymerase II fifth largest subunit, E
At4g17695	0.83	0.005	RNA	regulation of transcription.G2-like transcription factor family, GARP	KAN3__Homeodomain-like superfamily protein
At1g77920	-0.84	0.000	RNA	regulation of transcription.bZIP transcription factor family	TGA7__bZIP transcription factor family protein
At4g16610	0.75	0.014	RNA	regulation of transcription.C2H2 zinc finger family	C2H2-like zinc finger protein
At1g75520	-0.71	0.005	RNA	regulation of transcription.C2H2 zinc finger family	SRS5__SHI-related sequence 5
At1g70790	-0.74	0.003	RNA	regulation of transcription.C2H2 zinc finger family	Calcium-dependent lipid-binding (CaLB domain) family protein
At5g13920	-0.81	0.001	RNA	regulation of transcription.C2H2 zinc finger family	GRF zinc finger / Zinc knuckle protein
At1g32700	-0.71	0.005	RNA	regulation of transcription.unclassified	PLATZ transcription factor family protein
At5g10140	0.84	0.004	RNA	regulation of transcription.MADS box transcription factor family	AGL25_FLC_FLF_RSB6__K-box region and MADS-box transcription factor family protein
At4g01060	-0.71	0.005	RNA	regulation of transcription.MYB-related transcription factor family	CPL3_ETC3__CAPRICE-like MYB3
At5g16600	-0.74	0.003	RNA	regulation of transcription.MYB domain transcription factor family	AtMYB43_MYB43__myb domain protein 43
At1g16490	-0.88	0.000	RNA	regulation of transcription.MYB domain transcription factor family	ATMYB58_MYB58__myb domain protein 58
At3g07260	-0.78	0.002	RNA	regulation of transcription.FHA transcription factor	SMAD/FHA domain-containing protein
At5g06510	-0.72	0.005	RNA	regulation of transcription.CCAAT box binding factor family, HAP2	NF-YA10__nuclear factor Y, subunit A10
At4g39780	0.74	0.016	RNA	regulation of transcription.APETALA2/Ethylene-responsive element binding protein family	Integrase-type DNA-binding superfamily protein
At1g14340	-0.73	0.004	RNA	RNA binding	RNA-binding (RRM/RBD/RNP motifs) family protein

Table 7.2 List of genes having their expression correlating with penetrance of *amiPPD* in wild type. Only genes involved in transport, signaling, and RNA MapMan categories as well as direct target of PPD2 are shown. The expression levels of the underlined genes is shown in Figure 7.7D. *; direct target of PPD2, .

AT code	penetrance	Correlation	adjP-value		MapMan category	Gene description
At1g60960	Rosette	0.74	0.002		metal	ATIRT3_IRT3__iron regulated transporter 3
At3g27020*	Rosette	0.74	0.002	Transport	peptides and oligopeptides	AtYSL6_YSL6__YELLOW STRIPE like 6
At3g05030	Rosette	0.73	0.002	Transport	unspecified cations	ATNHX2_NHX2__sodium hydrogen exchanger 2
At1g65410	Rosette	-0.77	0.001	Transport	ABC transporters and multidrug resistance systems	ABCI13_ATNAP11_NAP11_TGD3__non-intrinsic ABC protein 11
At3g04090	Rosette	-0.72	0.003	Transport	Major Intrinsic Proteins.SIP	SIP1;1_SIP1A__small and basic intrinsic protein 1A
At1g80300	Rosette	-0.75	0.002		misc	ATNTT1_NTT1__nucleotide transporter 1
At5g54960*	Rosette	0.73	0.002		fermentation.PDC	PDC2__pyruvate decarboxylase-2
At1g28440	Rosette	0.75	0.001		receptor kinases.leucine rich repeat XI	HSL1__HAESA-like 1
At2g46700	Rosette	0.77	0.001		calcium	ATCRK3_CRK3__CDPK-related kinase 3
At2g15900	Rosette	0.72	0.002	Signaling	phosphinositides	Phox-associated domain;Phox-like;Sorting nexin, C-terminal
At2g44690	Rosette	-0.72	0.003	Signaling	G-proteins	ARAC9_ATROP8_ROP8__Arabidopsis RAC-like 9
At1g52380	Rosette	-0.79	0.001	Signaling	G-proteins	NUP50 (Nucleoporin 50 kDa) protein
At5g19010	Rosette	0.74	0.002		MAP kinases	MPK16__mitogen-activated protein kinase 16
At1g18160	Rosette	0.71	0.003		MAP kinases	Protein kinase superfamily protein
At1g43190	Rosette	-0.73	0.002		processing	PTB3__polypyrimidine tract-binding protein 3
At3g20390	Rosette	-0.75	0.002		processing.ribonucleases	endoribonuclease L-PSP family protein
At5g01310	Rosette	0.72	0.002		regulation of transcription.bHLH,Basic Helix-Loop-Helix family	APTX__APRATAXIN-like
At1g47870	Rosette	-0.72	0.003		regulation of transcription.E2F/DP transcription factor family	ATE2F2_ATE2FC_E2FC__winged-helix DNA-binding transcription factor family protein
At4g04890	Rosette	0.73	0.002		regulation of transcription.HB,Homeobox transcription factor family	PDF2__protodermal factor 2
At4g24630	Rosette	0.75	0.001	RNA	regulation of transcription.unclassified	DHHC-type zinc finger family protein
At4g21430	Rosette	0.71	0.003	RNA	regulation of transcription.unclassified	B160__Zinc finger, RING-type;Transcription factor jumonji/aspartyl beta-hydroxylase
At2g45660	Rosette	0.71	0.003		regulation of transcription.MADS box transcription factor family	AGL20_ATSOC1_SOC1__AGAMOUS-like 20
At3g10640	Rosette	-0.72	0.003		regulation of transcription.SNF7	VPS60.1__SNF7 family protein
At5g13680	Rosette	0.77	0.001		regulation of transcription.Histone acetyltransferases	ABO1_AtELP1_ELO2__IKI3 family protein
At5g47790	Rosette	0.79	0.001		regulation of transcription.FHA transcription factor	SMAD/FHA domain-containing protein
At1g34220*	Rosette	0.77	0.001		unknown	Regulator of Vps4 activity in the MVB pathway protein
At1g34130*	Seletive	0.71	0.003		stress.abiotic.drought/salt	STT3B__staurosporin and temperature sensitive 3-like b
At2g36880*	Seletive	-0.71	0.003		amino acid metabolism.synthesis.aspartate family.methionine	MAT3__methionine adenosyltransferase 3

Extended Data Table 7.1 Phenotype of *DN-DA1*^{OE} transgenics in 17 Arabidopsis accessions per independent transgenic line. Heat maps representing, per independent transgenic line, the average percent difference in each leaf area between *DN-DA1*^{OE} transgenics and their corresponding wild type (left) and the estimated expressivity (middle) of *DN-DA1*^{OE}. More than 10 plants were analyzed. Fold changes of expression levels of *DA1* compared to wild type analyzed by qPCR in T_2 generation are shown on the right panel (L1-L5; independent transgenic lines).

Accession	Lines	cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15	Sel	Ros	DA1
ICE61	1	84	116	111	130	114	123	95	65	58	42	35	39	34	48		119	91	3.11
	2	110	121	111	117	112	115	113	99	95	77	86	92	79	96		116	104	2.21
	3	94	125	116	122	117	125	100	81	66	49	49	50	45	60		121	96	5.09
	4	113	108	103	104	101	107	98	98	91	83	86	90	74	97		105	99	1.86
An-1	1	81	93	96	94	85	98	96									94	92	5.3
	2	79	108	97	100	99	100	84									102	110	4.9
	3	86	114	115	125	113	111	74									118	106	6.7
	4	95	116	103	106	104	105	83									108	103	8.3
	5	107	142	142	151	142	130	100									145	132	6.0
Cvi-0	1	104	140	122	117	120	100	106	107	72							126	114	5.35
	2	96	128	101	110	99	90	71	72	52							113	102	7.57
Ler-0	1	116	152	142	149	122	101	70	61								147	131	8.4
	2	98	121	111	124	107	96	77	54								119	112	7.1
	3	95	139	113	134	121	103	83	76								129	118	6.0
	4	97	127	119	127	113	95	75	52								124	117	9.4
Blh-1	1	118	170	139	152	123	115	102	82	76	63	71					154	121	7.5
	2	90	125	123	126	113	102	86	67	55	43	38					125	103	6.0
	3	107	157	147	146	123	111	90	71	70	54	48					150	117	4.7
	4	84	133	128	142	119	116	89	78	78	51	42					134	108	8.4
ICE138	1	128	135	163	158	156	120	76	88	77	56	57	53				152	121	5.9
	2	110	115	102	104	115	132	122	139	142	147	173	166				107	120	3.7
ICE97	1	96	131	125	133	132	126	115	122	101	92	94	83	92	71	75	130	115	3.7
Ey15-2	1	77	102	107	107	103	107	102	103	102	103	105	105				105	101	2.7
	2	80	95	81	84	70	74	68	63	54	48	39	34				86	66	5.5
	3	73	109	98	111	92	85	69	60	48	41	38	34				106	77	3.9
	4	77	118	125	130	123	116	94	80	67	43	39	38				124	92	8.8
C24	1	103	128	136	144	141	126	114	105	88	61	52	37	32			136	98	4.2
	2	74	99	119	132	133	109	97	72	47	29	22	14	13			117	74	9.8
	3	92	120	106	104	87	51	38	17	10	7	4	4	3			110	43	9.6
Yeg-1	1	96	116	116	109	108	97	102	107	106							114	112	2.5
WalhaesB4	1	75	166	146	153	132	130	111	95	88	79	77	81	76	102		155	114	10.2
	2	78	128	127	130	108	118	118	105	105	102	99	98	93	85		128	111	12.7
	3	69	108	97	95	94	107	108	113	114	105	120	117	138	108		100	104	4.2
Sha	1	92	117	128	133	111	107	83	74	86	65	93					126	109	3.8
	2	86	125	143	138	117	105	80	57	38	23	31					135	108	4.2
	3	100	129	135	133	115	113	82	66	51	35	30					132	108	2.6
	4	93	119	130	132	104	91	69	41	34	27	32					127	97	4.0
	5	90	131	140	133	109	100	85	48	53	43	44					135	105	2.8
Col-0	1	92	116	107	100	105	98	86	77	71	66	60	55	49			107	90	6.8
	2	91	116	105	105	108	98	85	77	71	62	51	57	59			108	91	10.2
	3	99	113	101	99	97	93	77	76	65	61	56	47	45			104	86	6.1
	4	93	105	83	85	78	82	73	64	63	58	46	39	27			91	74	10.6
	5	99	124	108	112	107	103	90	84	82	78	64	65	70			115	100	10.9
ICE153	1	85	67	74	77	83	95	93	101	95	79	73	56	66	57		73	84	3.4
	2	113	136	138	144	132	131	104	106	104	78	76	67	74	72		139	121	3.2
ICE163	1	101	125	116	121	107	114	100	104	95	68						121	94	6.0
Oy-0	1	117	147	123	123	117	114	98	94	90	73	70	53	61	72	60	131	104	9.9
	2	102	141	128	132	130	130	123	114	121	95	83	92	79	95	98	134	116	12.4
	3	123	152	132	140	130	135	128	120	128	122	106	113	109	140	121	141	127	8.0
	4	108	132	128	124	126	121	115	107	107	98	89	88	89	98	111	128	116	8.2
ICE75	1	131	129	111	121	109	103	99	94	94	88	87	81	79	80	79	120	100	2.6
	2	129	127	112	116	106	105	102	97	95	93	100	100	90	89	84	118	98	4.1

Chapter 8

Comparative analysis of destructive and non-destructive leaf area measurement methods in various genotypes in Arabidopsis

Youn-Jeong Nam[1,2], Stijn Dhondt[1,2], Dorota Herman[1,2], Frederik Coppens[1,2], Twiggy Van Daele[1,2], Dirk Inzé[1,2*] & Nathalie Gonzalez[1,2*]

[1]Department of Plant Systems Biology, VIB B-9052 Gent, Belgium, [2]Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Gent, Belgium
* These authors contributed equally to this work.

Contributions: Y.J.N. was the main author of this work. Y.J.N., and T.V.D. conducted experimental work. S.D., D.H., F.C. and Y.J.N. analysed the data. D.I. and N.G. supervised the research. D.I. and N.G. contributed to the writing of this chapter.

Introduction

Plant growth is an important trait for agriculture since it is directly related to crop yield. Thus, it is valuable to dissect this multi-factorial trait and understand its regulation. Growth can be described in many ways depending on which parameters are measured (weight, area, length) and at which level (plant, leaf or cellular level) (Dhondt et al., 2013). Furthermore, since growth is a dynamic process, time-component parameters need to be assessed such as relative growth rate (RGR) providing a measure of the speed of plant growth (Dhondt et al., 2014).

In order to measure shoot growth, various methods, destructive and non-destructive, have been used at different levels, such as quantification of the above ground plant or organ size (Blomme et al., 2014; Vanhaeren et al., 2014). In Arabidopsis, at rosette level, destructive techniques correspond to weighting the biomass of the above ground part of the plant or a part of the plant or making leaf series to analyze individual leaf area. However, these techniques limit the amount of data that can be obtained since measurements can only be made at a small number of time points due to the time consuming and tedious effort needed. Since leaf growth is a complex dynamic trait, detailed information such as time-resolved data are required and this information can be obtained from non-destructive phenotyping systems. The relatively flat shape of the Arabidopsis rosette allows for measuring projected rosette area with two dimensional digital imaging system. Nowadays, numerous automated image phenotyping systems (non-destructive) have been developed to measure projected rosette leaf area in Arabidopsis by using robotic platforms to follow the dynamics of rosette size in soil (Granier et al., 2006; Walter et al., 2007; Jansen et al., 2009; Arvidsson et al., 2011; Skirycz et al., 2011; De Vylder et al., 2012; Zhang et al., 2012; Tisné et al., 2013; Apelt et al., 2015) and also *in vitro* (Dhondt et al., 2014). Projected rosette area defined from top-view images can be linked to leaf function since this area is corresponding to the area where photosynthesis occurs. The imaging analysis of the rosette provides data sets in a non-destructive manner, detailed time-resolution, and saves time compared to tedious job such as making leaf series. However, phenotyping based on non-destructive analysis might also have limitations for measuring leaf growth. The rosette area obtained from a top-view image can be influenced by different factors such as leaf shape and overlap of the leaves during growth or diurnal movement.

Here, we performed a comparative analysis of the phenotype of different transgenic lines in 17 natural Arabidopsis accessions based on non-destructive data set and leaf series data (destructive). We used an in-house phenotyping system, MIRGIS (multi-camera *in vivo* rosette growth imaging system), which allows to grow Arabidopsis plants in an optimal environment and take pictures of the rosette daily for non-destructive data analysis. After imaging the plants till 25 DAS, rosette leaves were harvested to make leaf series and measure individual leaf area. Total rosette area was calculated based on these leaf series data. By comparing these measurements with the projected rosette area based on imaging, we show that there might be certain information incorrectly interpreted or missed by using only non-destructive data analysis, although also interesting information can be extracted from the images. We highlight the need to adapt the phenotyping methods in function of the specificity of each accession and each genotype.

Results and discussion

Images of *GA20ox1^{OE}, amiPPD*, and *DN-DA1^{OE}* transgenic lines from MIRGIS

In order to analyze the expressivity of transgenes in various genotypes, the effect of altering the expression of three growth promoting genes- *GA20ox1* (see chapter 5)*, PPD* (see chapter 6), and *DA1* (see chapter 7) - was analyzed in 17 Arabidopsis accessions. The 17 accessions originate from different regions around the world (Table 4.1) and show a wide range of leaf growth-related phenotypes (see chapter 5) as well as different flowering times. The three genes are known to positively regulate leaf growth when their expression is altered in a Col-0 background. Transgenic plants overexpressing *GA20ox1* display longer and larger leaves, longer petioles compared to wild type and flower earlier (Huang et al., 1998). Down-regulation *PPD* by overexpression of an artificial miRNA, *amiPPD*, causes the formation of larger leaves with a dome-shape phenotype (Gonzalez et al., 2015). A transgenic line overexpressing a dominant negative form of DA1, *DN-DA1 ^{OE},* produces larger leaves with rounder shape compared to control plants (Li et al., 2008).

Generally, from the non-destructive images various rosette size related parameters can be extracted. For example, projected rosette area (PRA), convex hull, and compactness can be measured. Projected rosette area represents the area that is occupied by the rosette in a top-view image. Rosette convex hull is the area corresponding to the smallest convex set containing the rosette (Figure 8.1A). Compactness, also referred to as surface coverage, is defined by the ratio between the projected rosette area and the convex hull.

The *GA20ox1^{OE}, amiPPD*, and *DN-DA1^{OE}* transgenic lines in the 17 accessions were grown for 25 DAS in the MIRGIS phenotyping platform. To examine the shape and size of the rosette for the different genotypes generated, we compared rosette images and projected rosette areas obtained from the MIRGIS. As examples, four accessions per transgene are selected having a range of phenotypes from no changes in leaf shape and/or size to an obvious alteration and pictures of wild types and transgenics and their projected rosette area are shown in Figure 8.1. For *GA20ox1^{OE}*, most of the transgenic lines produced larger leaves with longer petioles leading to a bigger rosette convex hull (data not shown). However, some transgenic lines showed decreased rosette size such as in C24 and Ler-0. Furthermore, some transgenic lines produced narrower leaves than wild type rolling down at the edge, for example, in ICE163. Transgenic lines of *amiPPD* showed dome-shape leaves in most of the accessions except few such as C24. This dome-shape leaf makes that the plant seems to have narrow leaves and decreased leaf area. We observed that this leaf shape alteration is not always linked to enlarged leaf area (see chapter 6) since some accessions having dome-shape leaves did not show increased leaf area such as ICE163 and Ey15-2. By overexpression of *DN-DA1,* most transgenic lines produced larger and rounder leaves with shorter petiole compared to wild type. This phenotype makes that the transgenic plants have increased compactness such as the transgenics of Oy-0.

In summary, rosette phenotype from non-destructive analysis of each transgene in 17 accessions showed various outcome with different size and shape.

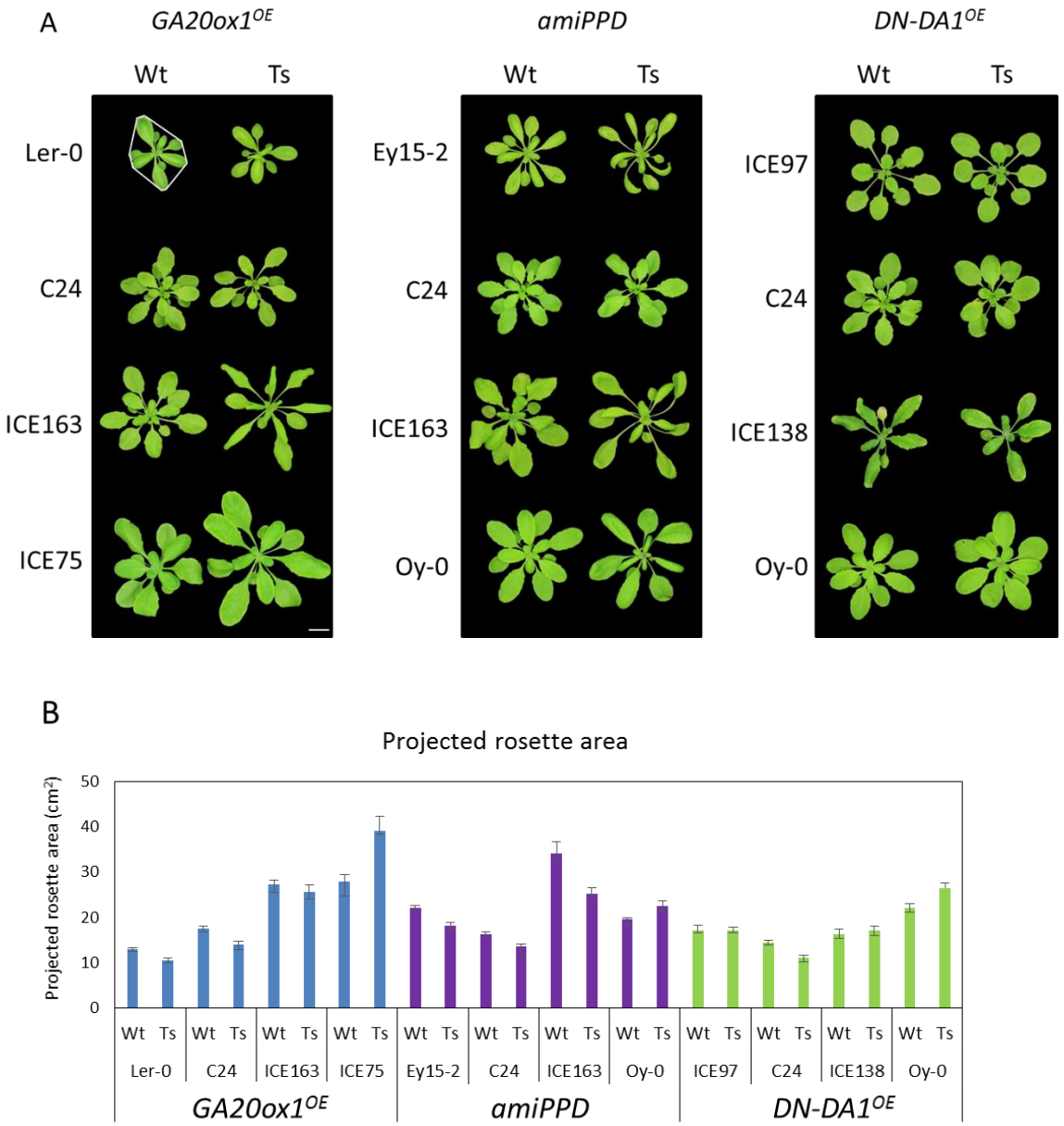


Figure 8.1 Phenotype of *GA20ox1*^{OE}, *amiPPD,* and *DN-DA1*^{OE} transgenics in different Arabidopsis accessions. (A) Image of 25-days-old rosette of representative transgenic plants and their corresponding wild types. An example of the convex hull is shown in the picture of wild-type plant of Ler-0 with white color line. Scale bar: 2 cm. (B) Projected rosette area of transgenic lines and their corresponding wild types shown in (A) at 25 DAS. At least 10 plants were analyzed for each genotype. Error bars represent standard errors.

Comparison of leaf size related parameters from different data sets

Next, we compared leaf size related phenotypic data produced by MIRGIS and from leaf series analysis (Table 8.1). From MIRGIS, projected rosette area (PRA) was obtained for each transgenic line and the corresponding wild type per accession from 4 DAS to 25 DAS, and the ratio to wild type was calculated (RPRA). The leaf series analysis (LS) was done at 25 DAS with same plants used for PRA calculation and the ratio to wild type of size of each individual leaf was also

calculated. 'Rosette' expressivity (Ros) was estimated from the leaf series data by calculating the ratio of total rosette area of the transgenic compared to its corresponding wild type.

In order to analyze how the phenotype of various genotypes can affect the accuracy of different leaf size related phenotypic data, we checked the consistency of these three different leaf size related phenotypes. The consistency was evaluated by comparing RPRA and Ros, and Ros and LS (Table 8.1). For example, if more than 75% of the transgenic lines showed an increased leaf area compared to wild type from LS as well as Ros, LS and Ros are evaluated as consistent. In these transgenic lines, however, if less than 75% of transgenic lines had small values of RPRA at 25 DAS, the Ros and RPRA are non-consistent.

In general, almost half of the accessions showed a consistent RPRA and Ros values for the three transgenes although *amiPPD* transgenic plants exhibited lower consistency (Table 8.1). This lower consistency is probably because of the dome-shape leaf phenotype in *amiPPD* transgenic lines making that the leaf seems narrower from a top-view. For example, in C24 in which *amiPPD* lines do not have dome-shape leaf phenotype, consistency between RPRA and Ros as well as between PRA and RA was found while in the Ey15-2 transgenic lines, having leaf shape alteration, the two values are less consistent compared to the other accessions (Figure 8.2A and Figure 8.2B). On the other hand, Ros and LS data in *amiPPD* transgenic lines are consistent in all accessions (Table 8.1). Also high consistency between Ros and LS was observed in *DN-DA1*^{OE} transgenic lines. We observed that low consistency between RPRA and Ros in transgenic lines of *DN-DA1*^{OE} was mainly found in early-flowering accessions (Table 8.1). In these accessions, for example in Ler-0, the overexpression of *DN-DA1*^{OE} led to delayed-flowering compared to wild type (Figure 8.3B). In contrast, in Oy-0 in which no inflorescence is observed at 25 DAS both in wild type and the transgenic lines consistency between RPRA and Ros data as well as between PRA and RA was found (Figure 8.3A). Thus, the PRA of Ler-0 wild type was calculated incorrectly because of cauline leaves that were considered as rosette. These data show that if plants flower at different time points, it has to be considered for automated analysis in order to obtain accurate results. For example, only early time points (before flowering) have to be analyzed to allow accurate estimation. For both comparisons, the consistency for rosette area related parameters in *GA20ox1*^{OE} transgenic lines was relatively lower than for the two other transgenes (Table 8.1). One possibility to explain this lower consistency would be the early flowering phenotype observed in the transgenic lines of *GA20ox1*^{OE} since most of the non-consistent data were observed in accessions that have this phenotype such as Ler-0 (Figure 8.4B). In Ler-0, *GA20ox1*^{OE} transgenic lines flower even earlier than the wild type and produce less leaves. For LS data to compare each individual leaf, only leaves which are present in wild type can be compared with transgenic ones. It means that there would be some leaves in the transgenic lines that are not taken into account for LS measurements. However, from Ros data, all the leaves were measured for total rosette area. Furthermore it could be that the flower was also considered as leaves from top-view image thereby PRA values might be estimated bigger than RA (Figure 8.4B). In ICE75, in which both wild type and transgenic lines do not develop an inflorescence at 25 DAS, consistency between RPRA and Ros was observed (Figure 8.4A). But, what makes that the data consistency between Ros and LS in transgenic lines of *DN-DA1*^{OE} are higher than data in *GA20ox1*^{OE}? One reason can be the

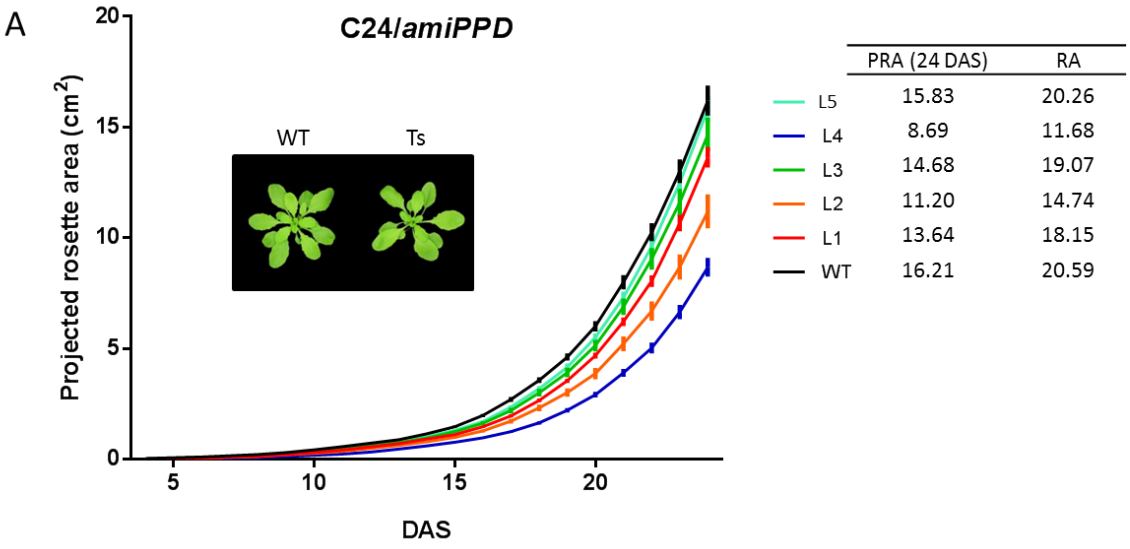
differences in the position of the leaves that were positively affected by either *GA20ox1*^{OE} or *DN-DA1*^{OE}. In transgenic lines of *GA20ox1*^{OE}, younger leaves were positively affected by the transgene while *DN-DA1* ^{OE} affects older leaves. Thus, while in transgenic lines of *GA20ox1*^{OE} some of the positively affected younger leaves become cauline leaves after flowering, transgenic lines of *DN-DA1* ^{OE} have the bigger older leaves. Although each leaf is bigger or similar in the *GA20ox1*^{OE} transgenic lines, the decreased total number of leaves compared to wild type made that total rosette area is reduced. However, transgenic lines of *DN-DA1*^{OE} showed bigger values in both individual leaf and total rosette area compared to wild type.

In summary, the data comparison between RPRA, Ros, and LS of the three transgenes introduced in the 17 accessions demonstrated that depending on the combination transgene/accession data consistency varies because of differences in variation of leaf number, leaf shape, and flowering time.

Table 8.1 Comparison between two different rosette area phenotyping results. Three different methods were used to estimate changes in rosette area: ratio of projected rosette area to wild type, 'rosette' expressivity (see Methods), leaf series data corresponding to the ratio to the wild type. These sets data were compared in two ways: ratio of projected rosette area vs 'rosette' expressivity and 'rosette' expressivity vs leaf series data. When more than 75% of transgenic lines shows consistency between two data sets, it is considered as consistent. RPRA; ratio of projected rosette area, Ros; 'rosette' expressivity, LS; leaf series, C; consistent, N; non-consistent. The percentage of score of consistency per transgene is shown. Early flowering; inflorescence appear before 25DAS, Medium flowering; inflorescence appear between 25 and 35 DAS, Late flowering inflorescence appear after 35 DAS.

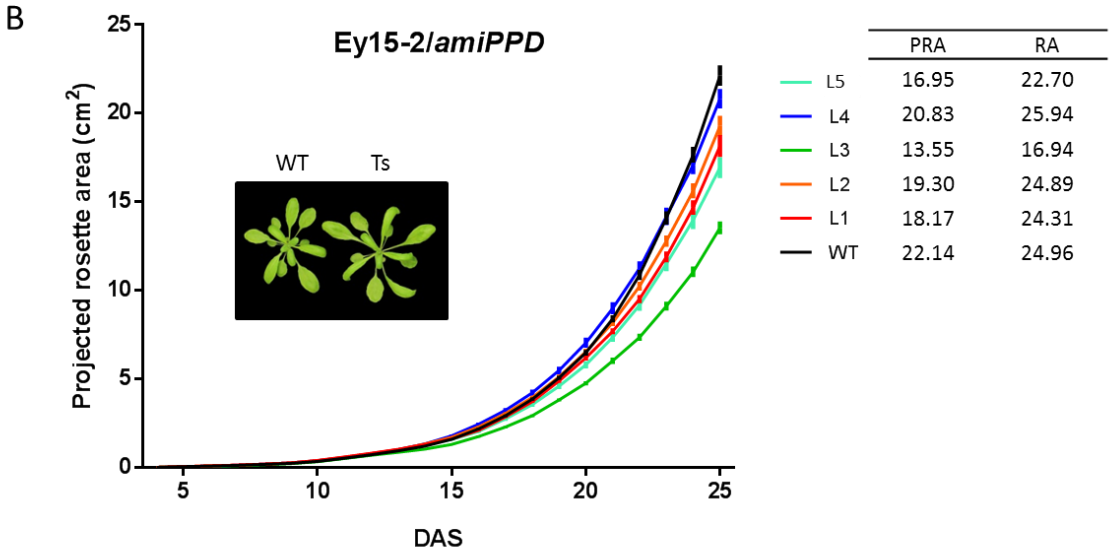
		GA20ox1^{OE}			*amiPPD*			*DN-DA1*^{OE}		
Accession	Flowering time	RPRA	Ros	LS	RPRA	Ros	LS	RPRA	Ros	LS
ICE61	L	C	C	C	C			C		C
An-1	E	N	C	C	C			N		C
Cvi-0	L	N	N	C	C			N		C
Ler-0	E	N	N	N	C			N		C
Blh-1	E	N	N	N	C			N		C
ICE138	L	N	C	C	C			N		N
ICE97	L	C	C	\	\			N		C
Ey15-2	M	C	N	N	C			C		C
C24	L	C	C	C	C			C		C
Yeg-1	L	C	C	N	C			C		C
WalhaesB4	M	C	C	C	C			C		C
Sha	M	C	C	N	C			C		C
Col-0	M	C	C	N	C			C		C
ICE153	L	N	C	C	C			C		C
ICE163	L	N	C	N	C			C		C
Oy-0	M	C	C	N	C			C		C
ICE75	L	C	C	\	\			N		C
Consistent %		59		76	47		100	59		94

\; no transgenics
E; Early flowering
M; Medium flowering
L; Late flowering



	PRA (24 DAS)	RA
L5	15.83	20.26
L4	8.69	11.68
L3	14.68	19.07
L2	11.20	14.74
L1	13.64	18.15
WT	16.21	20.59

Accession	Line	cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	Ros
C24	L1	94	112	118	120	122	118	108	97	86	62	48	39	37	29	89
	L2	84	101	115	108	111	106	91	85	56	40	32	22	26	17	76
	L3	99	106	107	111	111	103	107	97	85	80	65	56	50	44	92
	L4	64	82	100	102	105	93	79	51	26	16	10	6	5	5	57
	L5	102	104	114	121	120	117	123	115	94	85	67	57	49	46	99



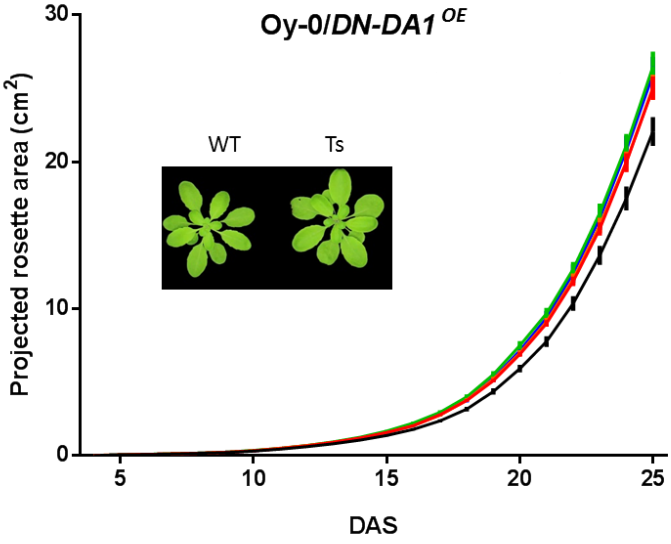
	PRA	RA
L5	16.95	22.70
L4	20.83	25.94
L3	13.55	16.94
L2	19.30	24.89
L1	18.17	24.31
WT	22.14	24.96

Accession	Lines	cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15	Ros
Ey15-2	L1	90	127	138	163	139	141	121	100	95	63	64	48	31	29	19	98
	L2	119	131	145	132	125	116	116	97	86	82	75	69	64	70	71	100
	L3	78	79	72	68	64	73	79	76	74	65	63	49	39	55	52	68
	L4	126	131	131	124	107	121	119	107	94	88	80	65	66	68	45	104
	L5	119	129	134	129	121	122	114	93	71	68	55	49	46	53	56	91

60 < | 60-80 | 80-90 | 90-95 | 95-105 | 105-110 | 110-120 | 120-140 | > 140

Figure 8.2 Comparison of three different rosette growth related datasets from *amiPPD* transgenics. Projected rosette area over time from 4 DAS to 25 DAS, leaf series data corresponding to the ratio to wild type at 25 DAS, and 'rosette' expressivity at 25 DAS, projected rosette area (PRA, at 24 DAS for C24) and rosette area (RA) at 25 DAS are shown with representative pictures of transgenics and wild type in C24 (A) and Ey15-2 (B). Red and green colors in leaf series data and 'rosette' expressivity (Ros) represent increased or decreased leaf area compared to wild type. At least 10 plants were analyzed for each genotype. Error bars represent standard errors.

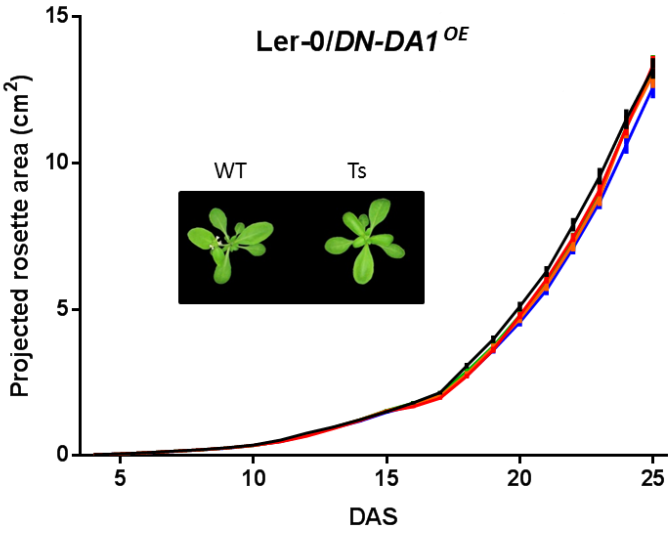
A



PRA	RA
25.95	30.01
26.51	32.88
25.10	30.15
25.03	27.08
22.08	25.94

Accession	Line	cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15	Ros
Oy-0	1	117	147	123	123	117	114	98	94	90	73	70	53	61	72	60	104
Oy-0	2	102	141	128	132	130	130	123	114	121	95	83	92	79	95	98	116
Oy-0	3	123	152	132	140	130	135	128	120	128	122	106	113	109	140	121	127
Oy-0	4	108	132	128	124	126	121	115	107	107	98	89	88	89	98	111	116

B



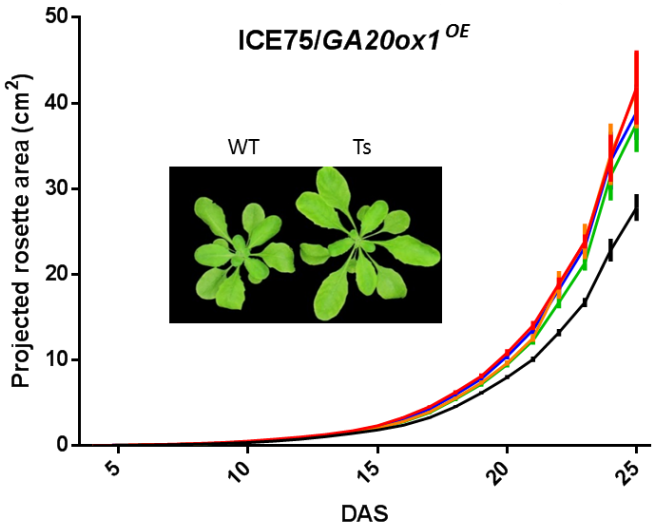
PRA	RA
12.61	15.69
13.26	15.09
13.04	15.36
13.36	17.53
13.25	14.07

Accession	Line	cot	L1/2	L3	L4	L5	L6	L7	L8	Ros
Ler-0	1	116	152	142	149	122	101	70	61	131
Ler-0	2	98	121	111	124	107	96	77	54	112
Ler-0	3	95	139	113	134	121	103	83	76	118
Ler-0	4	97	127	119	127	113	95	75	52	117

60< | 60-80 | 80-90 | 90-95 | 95-105 | 105-110 | 110-120 | 120-140 | > 140

Figure 8.3 Comparison of three different rosette area growth related datasets from *DN-DA1*^{OE} transgenics. Projected rosette area over time from 4 DAS to 25 DAS, leaf series data corresponding to the ratio to wild type at 25 DAS, 'rosette' expressivity at 25 DAS, projected rosette area (PRA) and rosette area (RA) at 25 DAS are shown with representative pictures of transgenics and wild type in Oy-0 (A) and Ler-0 (B). Red and green colors in leaf series data and 'rosette' expressivity (Ros) represent increased or decreased leaf area compared to wild type. At least 10 plants were analyzed for each genotype. Error bars represent standard errors.

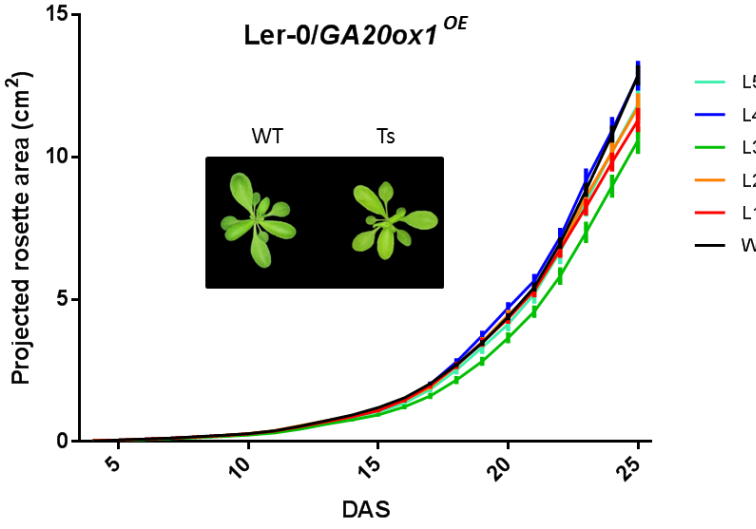
A



PRA	RA
38.95	38.40
37.62	36.57
41.51	36.45
41.86	41.53
27.82	28.82

accession	Line	Cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15	Ros
ICE75	L1	107	93	99	97	107	118	147	177	200	286	304	324	374	321	238	143
	L2	112	95	101	99	107	114	135	158	167	200	192	192	187	152	153	127
	L3	104	91	97	94	104	110	136	160	170	222	197	214	208	181	222	127
	L4	106	99	102	97	110	111	144	163	174	220	202	243	224	218	195	133

B



PRA	RA
11.86	10.91
12.87	12.26
10.60	10.77
11.77	10.67
11.32	10.33
12.89	14.64

accession	Line	Cot	L1/2	L3	L4	L5	L6	Ros
Ler-0	L1	82	80	77	77	90	96	64
	L2	87	85	90	94	111	134	67
	L3	62	66	67	64	78	87	72
	L4	81	92	86	91	107	126	81
	L5	78	81	86	89	99	113	74

60 < | 60-80 | 80-90 | 90-95 | 95-105 | 105-110 | 110-120 | 120-140 | > 140

Figure 8.4 Comparison of three different rosette growth related datasets from *GA20ox1*^{OE} transgenics. Projected rosette area over time from 4 DAS to 25 DAS, leaf series data corresponding to the ratio to wild type at 25 DAS, 'rosette' expressivity at 25 DAS, projected rosette area (PRA) and rosette area (RA) at 25 DAS are shown with representative pictures of transgenics and wild type in ICE75 (A) and Ler-0 (B). Red and green colors in leaf series data and 'rosette' expressivity (Ros) represent increased or decreased leaf area compared to wild type. At least 10 plants were analyzed for each genotype. Error bars represent standard errors.

Data divergence of non-destructive image analysis depending on the time point

Since non-destructive image analysis was done at numerous time points, we examined how the time point of non-destructive image analysis affects the interpretation of rosette size. Ratio of PRA from transgenics and wild type at two different time points (21 and 25 DAS) were compared to 'rosette' expressivity measured at 25 DAS (Figure 8.5). Based on Ros values, *GA20ox1*^{OE} transgenic lines of ICE75 produced 43% larger rosette leaves at 25 DAS. RPRA at different time points showed quite similar increase values compared to wild type. At 21 DAS, RPRA of this transgenic line is 39% larger than wild type and 4 days after, it showed 50% increase (Figure 8.5A). In ICE75, the data therefore showed consistency. In the case of ICE61/*GA20ox1*^{OE}, almost no change in RPRA was found between the two time points. But a big difference was found between RPRA and the actual rosette area since the real value was 15% larger than the estimated one (Figure 8.5B). This difference might be due to the leaf shape alteration in the transgenic lines having wavy and therefore narrower leaves compared to wild type probably leading to an underestimation of the area from the images (Figure 8.5B). Interestingly, one of the *GA20ox1*^{OE} transgenic lines of An-1 showed dramatic differences between RPRA at 21 DAS, RPRA at 25 DAS, and Ros at 25 DAS (Figure 8.5C). At 21 DAS, the transgenic lines had 5% increase in RPRA and the RPRA became 64% bigger than wild type at 25 DAS. However, this increase is a wrong interpretation because of the presence of an inflorescence that was considered as part of the rosette from the non-destructive image analysis (transgenic plants started to bolt already at 19 DAS). In fact the real rosette area at 25 DAS was 51% decreased compared to wild type. In the representative picture of An-1/*GA20ox1*^{OE}, the transgenic plant showed long inflorescences with flowers (Figure 8.5C). Therefore, for the automated non-destructive image analysis, the data accuracy can vary depending on not only the genotype but also the time point analyzed. For example, flowering time in comparison with wild type can influence the accuracy of the data, thus, by correcting the automated non-destructive image data sets, incorrect interpretations can be avoided.

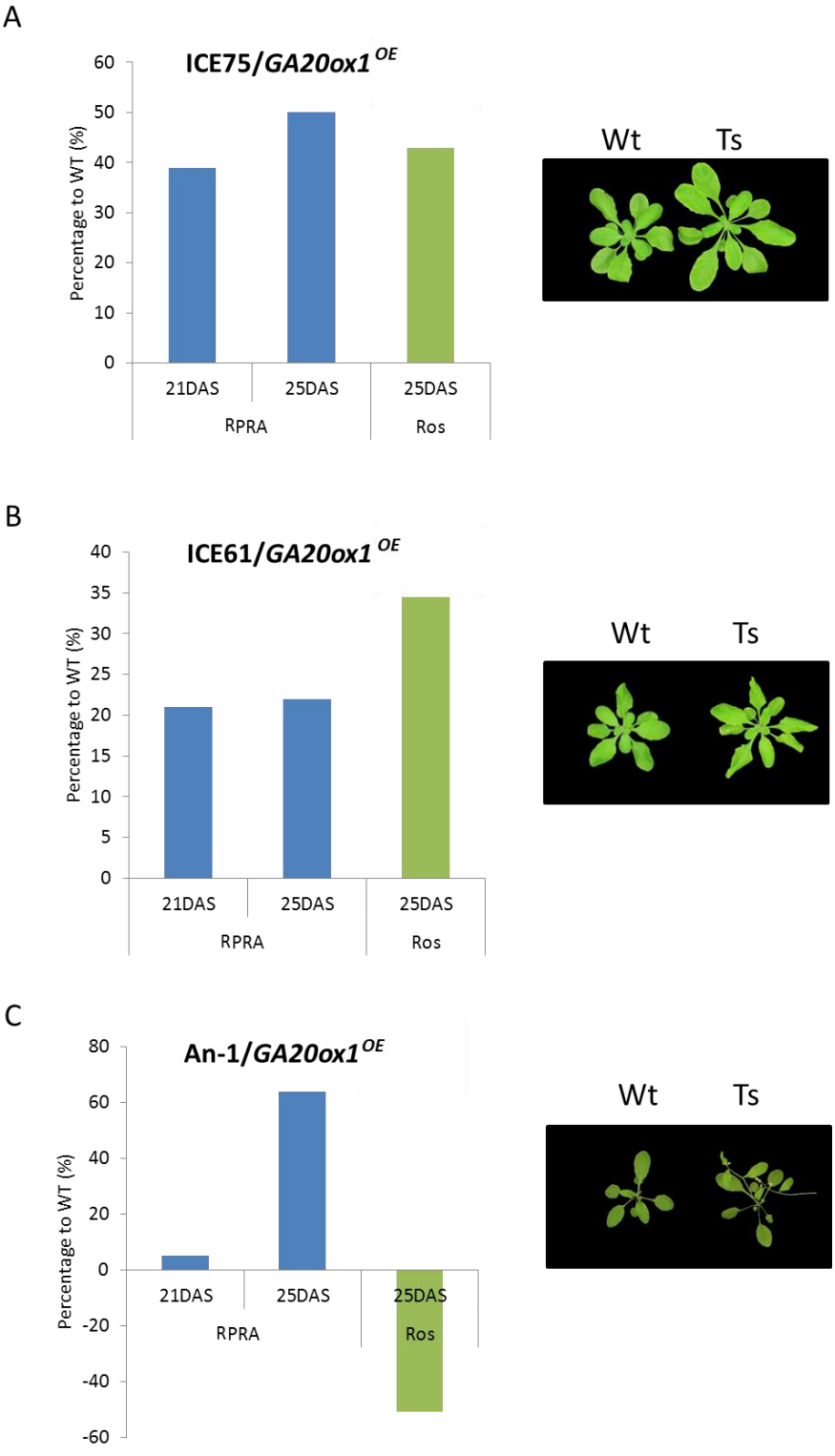


Figure 8.5 Comparison analysis between ratio to wild type of projected rosette area (RPRA) at different time points and percentage to wild type of real rosette area based on leaf series (Ros) in transgenic lines. The projected rosette area at two different time points (21DAS and 25 DAS, blue color bars) and rosette area at 25 DAS (green color bar) are shown with representative pictures of *GA20ox1*^{OE} transgenic plants and wild type in ICE75 (A), ICE61 (B), and An-1 (C). Wt; wild type, Ts; transgenics.

Analysis of compactness in the transgenic lines

Based on the MIRGIS data, we analyzed the compactness of the transgenic plants (Table 8.2 and Figure 8.6). We observed that the majority of the transgenic lines of *GA20ox1*^{OE} and *amiPPD* showed reduced compactness compared to wild type whereas *DN-DA1*^{OE} displayed increased compactness (Figure 8.6). In total, 88% of the accessions showed a clear difference in compactness between wild type and *GA20ox1*^{OE} lines and transgenic lines for 10 of these accessions displayed differences before 17 DAS (Table 8.2). Eighty percent of the *amiPPD* transgenics had a decreased compactness compared to wild type. *amiPPD* transgenic lines tend to show this difference of compactness earlier since most transgenics already have a less compact rosette at 14 DAS. For the transgenic plants of *DN-DA1*^{OE}, in 53% of the accessions, we found increased compactness compared to wild type which was visible before 17 DAS for 7 accessions.

Compactness can be influenced by several factors. For example, plants having many leaves with short petioles will have high value of compactness. Also leaf shape can affect compactness as leaves with rounder shapes can lead to high surface coverage.

To analyze the compactness during leaf growth, we compared its evolution in 3 accessions, C24, Ler-0, and ICE163, for the three transgenes from 4 DAS to 25 DAS (Figure 8.7, Figure 8.8, and Figure 8.9). For *GA20ox1*^{OE} transgenics, different changes in compactness overtime were found for the three accessions. In C24 accession, compactness in the transgenic plants became different at 21 DAS. This was probably due to the lower number of leaves in the transgenic plants leading to a smaller surface coverage (Figure 8.7A). However, transgenic lines of Ler-0 did not show alterations in compactness compared to wild type until 25 DAS. This might be to the absence of changes in petiole length and leaf size in the transgenic plants compared to wild type (Figure 8.7B). We observed that all transgenic lines of ICE163 have a less compact rosette than wild type at 16 DAS (Figure 8.7C). It seems that both increased petiole and leaf length and alteration of leaf shape leads to this decrease.

In the case of *amiPPD*, C24 transgenic plants also produce less leaves leading to reduced compactness compared to wild type at 21 DAS (Figure 8.8A). The *amiPPD* transgenic plants in Ler-0 showed reduced compactness already at 14 DAS compared to wild type (Figure 8.8B). Both increased petiole and leaf length and alteration of leaf shape might cause this decrease. Also *amiPPD* transgenic plants in ICE163 had decreased compactness at 14 DAS (Figure 8.8C). Mainly leaf shape alteration (dome-shape leaf) was observed in these transgenic and could explain the change in compactness.

DN-DA1^{OE} transgenics in C24 showed increased compactness compared to wild type at 11 DAS, but the compactness of wild type became similar to two of the transgenic plants at 25 DAS (Figure 8.9A). It might be that although size increases in older leaves of the transgenic plants causing higher compactness at early time point, wild type had more leaves that can increase the surface coverage at late time points. In Ler-0/*DN-DA1*^{OE}, no difference in compactness was observed (Figure 8.9B). A transgenic line of ICE163/*DN-DA1*^{OE} had increased compactness compared to wild type at 10 DAS (Figure 8.9C). The alteration of leaf shape (rounder) and increased area of older leaves caused overlap of the leaves which resulted in higher compactness

in the transgenic line. Interestingly, during leaf development, compactness of *DN-DA1*^{OE} rosettes in the three accessions decreased with 1 or 2 days of delay compared to the WT. Rosette compactness decreases as plant gets older because more leaves appear and the leaves start to expand (Dhondt et al., 2014). The delayed decrease of compactness might indicate slower speed of leaf development in the transgenic plants of *DN-DA1*^{OE}.

In conclusion, compactness varies greatly in function of the transgene and the genetic background and depends on the leaf shape, area and development stages of the plant.

Table 8.2 Overview of days after stratification (DAS) when the compactness of transgenics showed differences compared to wild type. The percentages represent the ratio of transgenics showing a difference in compactness towards their corresponding wild type.

Accession	*GA20ox1*^{OE}	*amiPPD*	*DN-DA1*^{OE}
ICE61	15	14	16
An-1	*	14	23
Cvi-0	23	*	13
Ler-0	*	14	*
Blh-1	22	14	15
ICE138	22	13	*
ICE97	12	\	*
Ey15-2	15	13	*
C24	21	21	11
Yeg-1	14	17	*
WalhaesB4	13	*	*
Sha	15	*	9
Col-0	15	14	*
ICE153	24	14	16
ICE163	16	14	10
Oy-0	15	14	23
ICE75	12	\	*
Percentage	88%	80%	53%

\; no transgenics
*; no difference with wild type

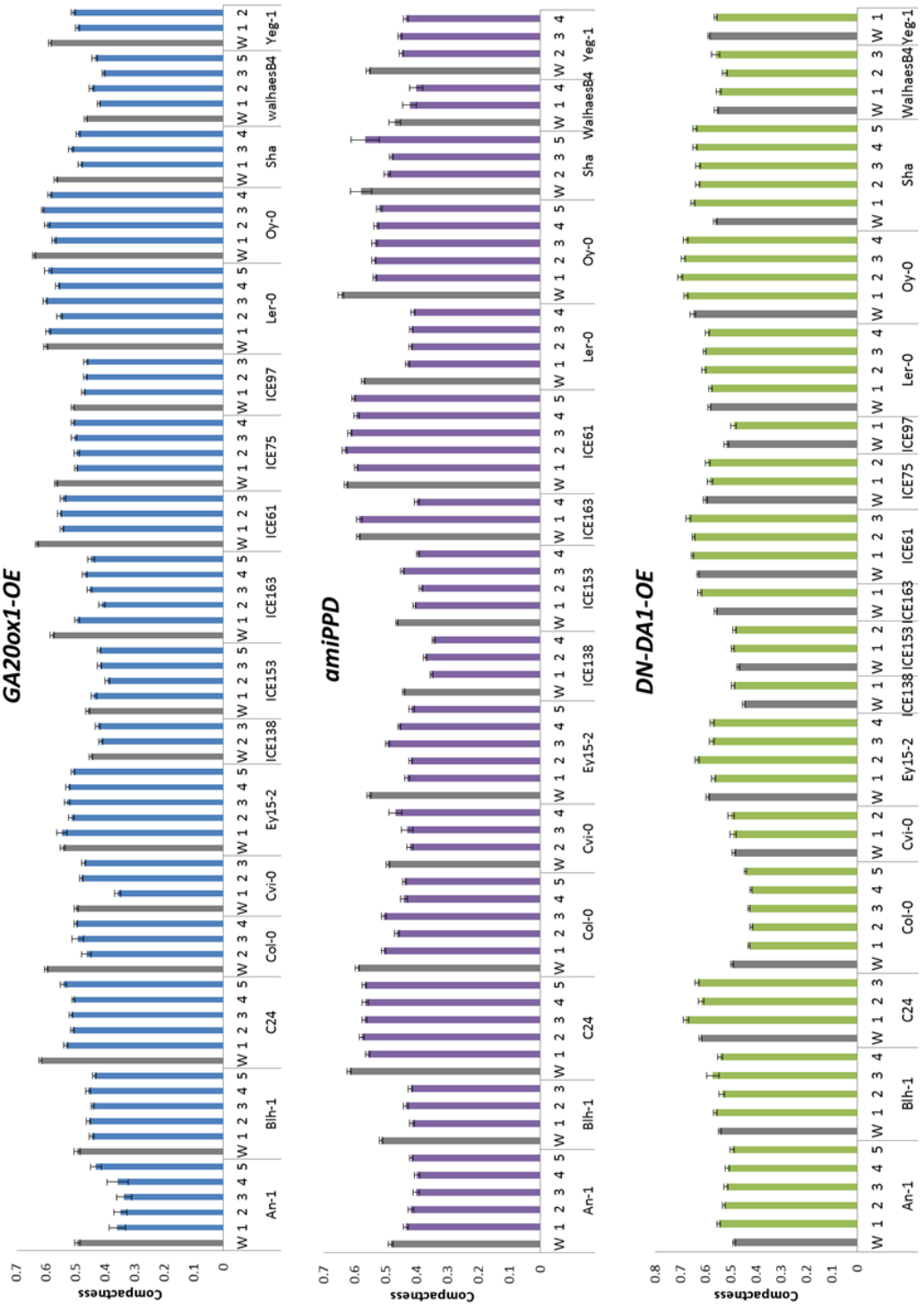
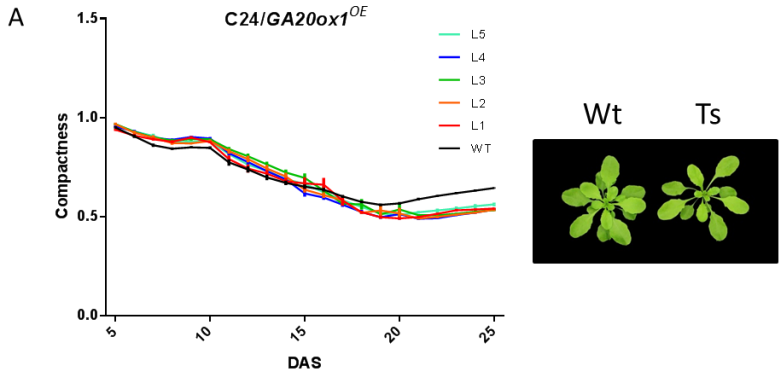
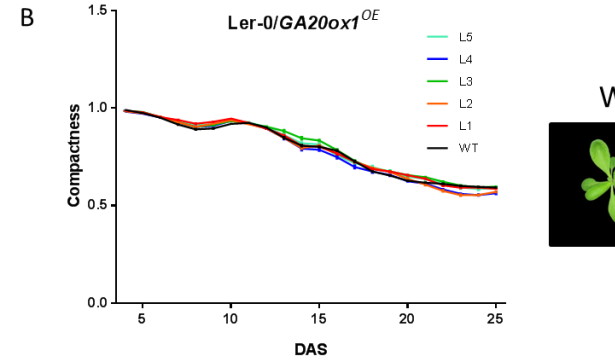


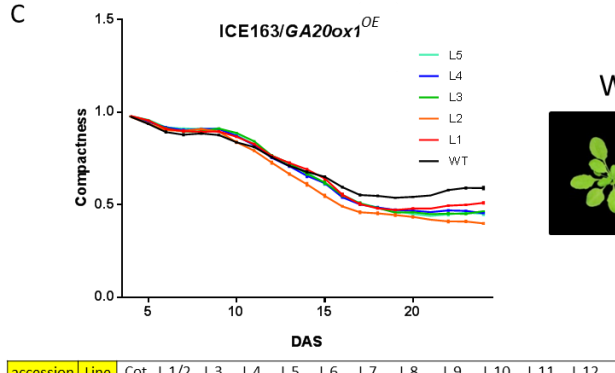
Figure 8.6 Compactness of GA20ox1OE, amiPPD, and DN-DA1OE transgenics in 17 Arabidopsis accessions. Compactness (projected rosette area/convex hull) of 23-days-old rosette of all transgenic plants and their corresponding wild types. At least 10 plants were analyzed in each genotype. Error bars represent standard errors.



accession	Line	Cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	Ros
C24	L1	82	87	92	97	107	98	104	102	109	83	70	65	52	44	91
	L2	77	78	89	95	95	88	86	81	67	54	44	31	26	22	72
	L3	66	77	90	92	94	97	94	91	80	61	58	36	28	23	77
	L4	69	77	91	98	105	100	95	90	83	70	61	42	34	37	81
	L5	82	94	93	100	104	102	98	89	83	69	52	38	32	35	82



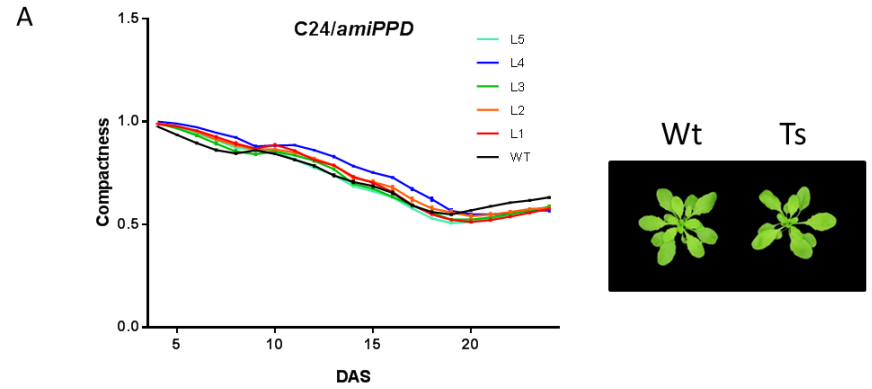
accession	Line	Cot	L1/2	L3	L4	L5	L6	Ros
Ler-0	L1	82	80	77	77	90	96	64
	L2	87	85	90	94	111	134	67
	L3	62	66	67	64	78	87	72
	L4	81	92	86	91	107	126	81
	L5	78	81	86	89	99	113	74



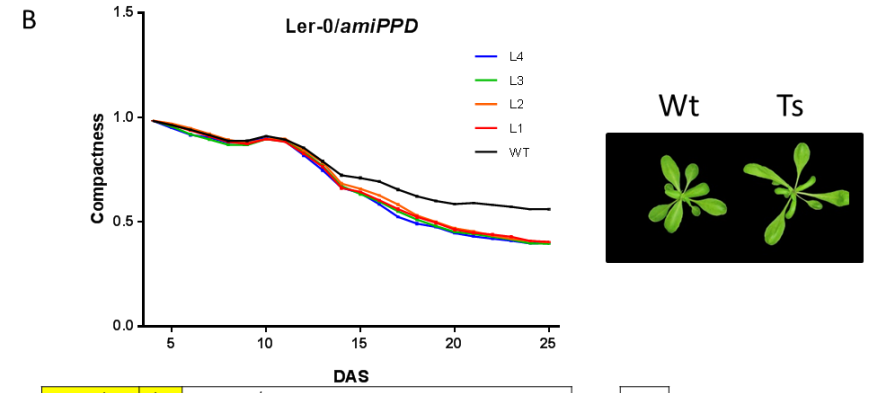
accession	Line	Cot	L 1/2	L 3	L 4	L 5	L 6	L 7	L 8	L 9	L 10	L 11	L 12	L 13	L 14	L 15	Ros
ICE163	L1	83	107	102	100	102	114	119	138	131	146	148	125	147	143	136	124
	L2	105	116	110	111	121	135	148	159	158	168	158	156	151	156	146	143
	L3	74	87	88	87	104	107	108	105	90	91	82	66	67	58	59	92
	L4	72	86	85	93	102	112	120	123	124	116	112	99	93	88	81	109
	L5	82	97	103	106	117	138	141	145	138	137	141	120	119	115	112	127

60 <	60-80	80-90	90-95	95-105	105-110	110-120	120-140	> 140

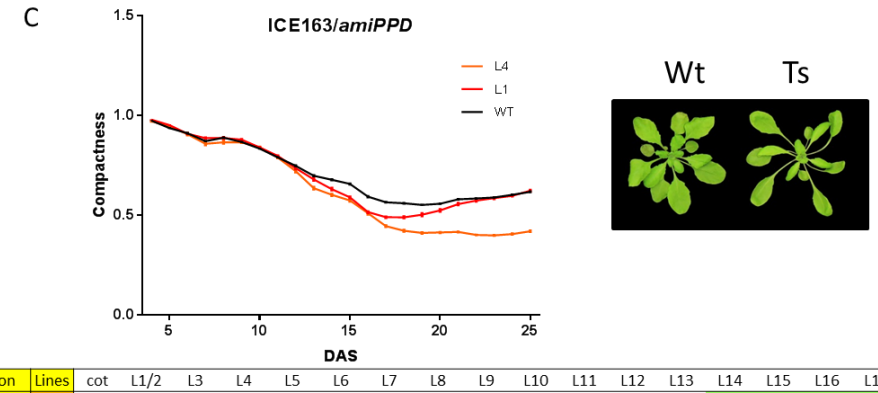
Figure 8.7 Compactness overtime, leaf series data and 'rosette' expressivity of *GA20ox1*^{OE} transgenics in three accessions. Compactness of the transgenic lines overexpressing *GA20ox1* in C24, Ler-0, and ICE163 from 4 DAS to 25 DAS. Representative pictures of the transgenics and wild-type plants are shown. Ratio to wild type of individual leaf area based on leaf series data and 'rosette' expressivity (Ros) are also shown. Red and green colors for leaf series data and 'rosette' expressivity (Ros) represent increased or decreased leaf area compared to wild type. At least 10 plants were analyzed in each genotype. Error bars represent standard errors.



Accession	Line	cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	Ros
C24	L1	94	112	118	120	122	118	108	97	86	62	48	39	37	29	89
	L2	84	101	115	108	111	106	91	85	56	40	32	22	26	17	76
	L3	99	106	107	111	111	103	107	97	85	80	65	56	50	44	92
	L4	64	82	100	102	105	93	79	51	26	16	10	6	5	5	57
	L5	102	104	114	121	120	117	123	115	94	85	67	57	49	46	99



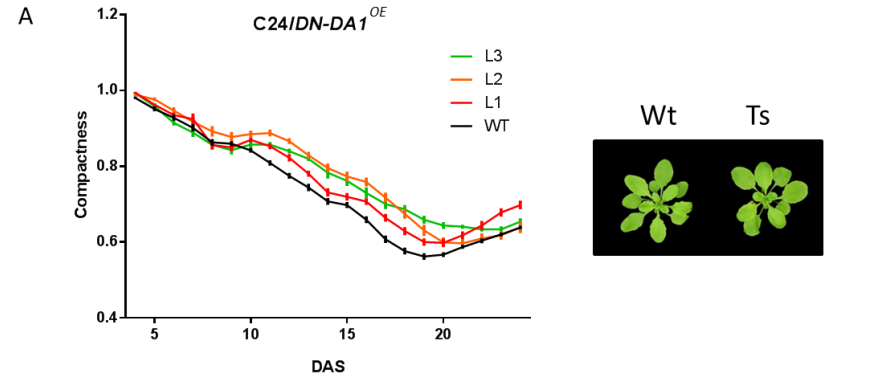
Accession	Line	cot	L1/2	L3	L4	L5	L6	L7	L8	Ros
Ler-0	L1	88	81	95	97	110	101	88	93	118
	L2	105	112	133	136	131	103	72	62	115
	L3	101	109	123	126	135	112	91	95	131
	L4	108	105	138	142	124	104	87	82	123



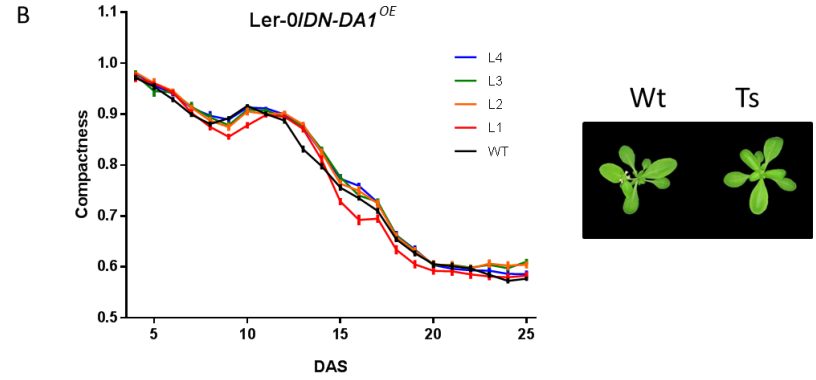
Accession	Lines	cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	L14	L15	L16	L17	L18	Ros
ICE163	L1	101	99	98	95	101	103	100	102	103	99	100	99	97	90	87	86	88	83	99
	L4	103	99	98	100	100	100	95	94	95	90	91	89	87	80	76	73	78	73	95

60 <	60-80	80-90	90-95	95-105	105-110	110-120	120-140	> 140

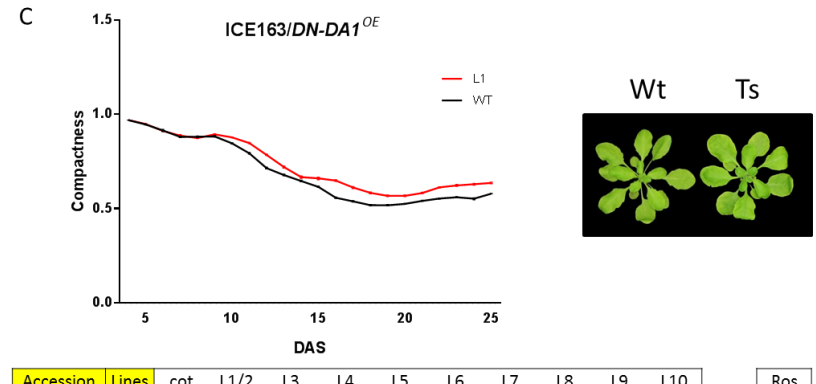
Figure 8.8 Compactness overtime, leaf series data and 'rosette' expressivity of *amiPPD* transgenics in three accessions. Compactness of transgenic lines of *amiPPD* in C24, Ler-0, and ICE163 from 4 DAS to 25 DAS. Representative pictures of transgenics and wild-type plants are shown. Ratio to wild type of individual leaf area based on leaf series data and 'rosette' expressivity (Ros) are also shown. Red and green colors in leaf series data and 'rosette' expressivity (Ros) represent increased or decreased leaf area compared to wild type. At least 10 plants were analyzed in each genotype. Error bars represent standard errors.



Accession	Line	cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12	L13	Ros
C24	1	103	128	136	144	141	126	114	105	88	61	52	37	32	98
C24	2	74	99	119	132	133	109	97	72	47	29	22	14	13	74
C24	3	92	120	106	104	87	51	38	17	10	7	4	4	3	43



Accession	Line	cot	L1/2	L3	L4	L5	L6	L7	L8	Ros
Ler-0	1	116	152	142	149	122	101	70	61	131
Ler-0	2	98	121	111	124	107	96	77	54	112
Ler-0	3	95	139	113	134	121	103	83	76	118
Ler-0	4	97	127	119	127	113	95	75	52	117



Accession	Lines	cot	L1/2	L3	L4	L5	L6	L7	L8	L9	L10	Ros
ICE163	1	101	125	116	121	107	114	100	104	95	68	94

60 < | 60-80 | 80-90 | 90-95 | 95-105 | 105-110 | 110-120 | 120-140 | > 140

Figure 8.9 Compactness overtime, leaf series data and 'rosette' expressivity of *DN-DA1*^{OE} transgenics in three accessions. Compactness of the transgenic lines overexpressing *DN-DA1* in C24, Ler-0, and ICE163 from 4 DAS to 25 DAS. Representative pictures of the transgenics and wild-type plants are shown. Ratio to wild type of individual leaf area based on leaf series data and 'rosette' expressivity (Ros) are also shown. Red and green colors in leaf series data and 'rosette' expressivity (Ros) represent increased or decreased leaf area compared to wild type. At least 10 plants were analyzed in each genotype. Error bars represent standard errors.

Conclusion

Non-destructive leaf area measurement methods provide detailed information such as time-resolved data which that destructive technique cannot generate, but it also has some limitations to estimate leaf area. By comparing projected rosette area with rosette area calculated from leaf series data, we showed that half of the transgenic lines in the different genetic backgrounds showed no-consistency between destructive and non-destructive data. This inconsistency was depending not only on the transgene but also on the genetic background. Interestingly, a lack of consistency was also observed between 'rosette' expressivity and leaf series data mainly because of a difference in the number of leaves between the transgenic and the corresponding wild type. Furthermore, we also found that the data accuracy of automated imaging analysis can vary depending on the time point of analysis. Therefore, this study suggested that in order to measure rosette area accurately by using automated imaging analysis, one has to consider several parameters to avoid incorrect interpretation caused by different genotypes, leaf shape, area and development stages of the plant. For example, if projected rosette area is not properly measured because of leaf shape changes, leaf series can be conducted to measure accurately this area. Also if plants are flowering early, this information should be taken into account for non-destructive image analysis.

Methods

Plant material and growth conditions

Seventeen Arabidopsis accessions were selected to cover a large genetic variation of *Arabidopsis thaliana* (Table 4.1) and used to generate three transgenic lines, *GA20ox1* overexpression (see chapter 5), *amiPPD* (see chapter 6), *DN-DA1* overexpression (see chapter 7). For the *in-vivo* growth analysis, five plants per pots were sown on soil. After one week, the seedling with a projected rosette area closest to median area of that genotype was selected per pot. While the growth of these seedlings was followed during further development, the other seedlings were removed. Average light intensity, supplied by cool-white fluorescent tubes (Spectralux Plus 36W/840; Radium), was around 100 mE m^{-2} s^{-1}. All plants were grown in soil under a 16-hours day/8-hours night regime at 21°C in MIRGIS.

Image acquisition, image processing and data analysis

For the *in-vivo* growth analysis, image acquisition was performed using Canon EOS 550D slr cameras equipped with a Canon EF 35mm f/2 objective. Pictures were automatically captured on a daily basis by a Perl script (www.perl.org) using the gPhoto2 library (www.gphoto.org). Image preprocessing and segmentation for the seedling selection and growth analysis was performed with C++ scripts using the OpenCV image analysis library (www.opencv.org). Parsing of quantitative measurements and further data analysis was performed with Perl scripts (www.perl.org). Graphs

of the calculated data were automatically plotted making use of the graphing utility gnuplot (www.gnuplot.info). More details on the image and data analysis procedures can be found elsewhere (Dhondt et al., 2014).

Leaf series data analysis

For the leaf series data, leaf area was log transformed to stabilize the variance. The mean model consisted of the main effects of *GA20ox1* overexpression on leaf size and their interaction term. Due to the unbalanced and complex nature of the data, the Kenward-Rogers approximation for computing the denominator degrees of freedom for the tests of fixed effects was used. An autoregressive structure was used to model the correlations between measurements done on the leaves originating from the same plant. The main interest was in the effect of the gene on leaf area for each leaf separately. Simple tests of effects were performed at each leaf between the transgenic lines and the corresponding wild type. Difference estimates were represented as % to the least-square means estimate of the wild type and leaf. Separate models were run for each accession as they were grown in separate experiments. For each experiment, the data was truncated so that there were at least two observations for each leaf of both the transgenic lines and the corresponding wild type. The analysis was performed with the mixed and plm procedure of SAS [Version 9.4 of the SAS System for windows 7 64bit. Copyright © 2002-2012 SAS Institute Inc. Cary, NC, USA (www.sas.com)]. Residual diagnostics were carefully examined.

'Rosette' expressivity was estimated as a ratio of the wild-type rosette area to that of a transgenic line. In case of 'rosette' expressivity per accession, the mean of 'rosette' expressivity per transgenic line of an accession has been taken.

References

Apelt, F., Breuer, D., Nikoloski, Z., Stitt, M., and Kragler, F. (2015). Phytotyping^{4D}: a light-field imaging system for non-invasive and accurate monitoring of spatio-temporal plant growth. Plant J. **82:** 693-706.

Arvidsson, S., Pérez-Rodriguez, P., and Mueller-Roeber, B. (2011). A growth phenotyping pipeline for *Arabidopsis thaliana* integrating image analysis and rosette area modeling for robust quantification of genotype effects. New Phytol. **191:** 895-907.

Blomme, J., Inze, D., and Gonzalez, N. (2014). The cell-cycle interactome: a source of growth regulators? J Exp Bot **65:** 2715-2730.

De Vylder, J., Vandenbussche, F., Hu, Y., Philips, W., and Van Der Straeten, D. (2012). Rosette Tracker: an open source image analysis tool for automatic quantification of genotype effects. Plant Physiol. **160:** 1149-1159.

Dhondt, S., Wuyts, N., and Inzé, D. (2013). Cell to whole-plant phenotyping: the best is yet to come. Trends Plant Sci. **18:** 428-439.

Dhondt, S., Gonzalez, N., Blomme, J., De Milde, L., Van Daele, T., Van Akoleyen, D., Storme, V., Coppens, F., Beemster, G.T.S., and Inzé, D. (2014). High-resolution time-resolved imaging of *in vitro* Arabidopsis rosette growth. Plant J. **80:** 172-184.

Gonzalez, N., et al. (2015). A Repressor Protein Complex Regulates Leaf Growth in Arabidopsis. Plant Cell **27:** 2273-2287.

Granier, C., et al. (2006). PHENOPSIS, an automated platform for reproducible phenotyping of plant responses to soil water deficit in *Arabidopsis thaliana* permitted the identification of an accession with low sensitivity to soil water deficit. New Phytol. **169:** 623-635.

Huang, S., Raman, A.S., Ream, J.E., Fujiwara, H., Cerny, R.E., and Brown, S.M. (1998). Overexpression of 20-oxidase confers a gibberellin-overproduction phenotype in Arabidopsis. Plant Physiol. **118:** 773-781.

Jansen, M., et al. (2009). Simultaneous phenotyping of leaf growth and chlorophyll fluorescence via GROWSCREEN FLUORO allows detection of stress tolerance in *Arabidopsis thaliana* and other rosette plants. Funct. Plant Biol. **36:** 902-914.

Li, Y., Zheng, L., Corke, F., Smith, C., and Bevan, M.W. (2008). Control of final seed and organ size by the *DA1* gene family in *Arabidopsis thaliana*. Genes Dev. **22:** 1331-1336.

Skirycz, A., et al. (2011). Survival and growth of *Arabidopsis* plants given limited water are not equal. Nat. Biotechnol. **29:** 212-214.

Tisné, S., et al. (2013). Phenoscope: an automated large-scale phenotyping platform offering high spatial homogeneity. Plant J. **74:** 534-544.

Vanhaeren, H., Gonzalez, N., Coppens, F., De Milde, L., Van Daele, T., Vermeersch, M., Eloy, N.B., Storme, V., and Inzé, D. (2014). Combining growth-promoting genes leads to positive epistasis in *Arabidopsis thaliana*. eLife **3:** e02252.

Walter, A., Scharr, H., Gilmer, F., Zierer, R., Nagel, K.A., Ernst, M., Wiese, A., Virnich, O., Christ, M.M., Uhlig, B., Jünger, S., and Schurr, U. (2007). Dynamics of seedling growth acclimation towards altered light conditions can be quantified via GROWSCREEN: a setup and procedure designed for rapid optical phenotyping of different plant species. New Phytol. **174:** 447-455.

Zhang, X., Hause Jr., R.J., and Borevitz, J.O. (2012). Natural genetic variation for growth and development revealed by high-throughput phenotyping in *Arabidopsis thaliana*. G3: Genes, Genomes, Genet. **2:** 29-34.

Part 3. General discussion and perspectives

General discussion and perspectives

World's demand for plant-derived products such as food, feed and bio-energy products is gradually increasing. To fulfill the high demand for agricultural products, genetically modified (GM) crops have been suggested as one of the possible ways to solve the problem. The first generation of GM crops focusing on herbicide tolerance and insect resistance are still dominating the current market for biotech crops. These traits are often independent and peripheral to the plant core metabolism and development and usually show a very high penetrance across different genotypes. For the next generation of GM crops, in order to further increase crop yield, gene engineering directly targeting the mechanisms responsible for growth would be needed. Since plant growth is a complex trait, these GM crops are more difficult to develop than the first generation GMs. Additionally, interactions between the inserted transgene and the genetic background might affect the expected phenotype and therefore creates an extra level of complexity. Different effects of a mutation in different genetic backgrounds have been observed in several organisms, including bacteria, yeast, worm, fruit fly, and mouse (reviewed in Chandler et al., 2013). However, the knowledge about how the genetic background affects the penetrance of a mutation in plants is poorly understood. Thus, in order to obtain stable effect of a transgene across different genetic backgrounds in crops, a deep mechanistic understanding of the effect of genetic perturbations in various genotypes is strongly required.

In this thesis, we analyzed the effect of transgenes involved in leaf growth regulation in different Arabidopsis accessions. The genes studied here, *GA20ox1*, *PPD2* and *DA1*, were previously shown as growth-promoting genes in Col-0 background (Huang et al., 1998; Li et al., 2008; Gonzalez et al., 2015). The three transgenes were introduced in 17 Arabidopsis natural accessions including Col-0, which are originating from around the world. The effect of these transgenes was analysed at morphological and molecular level. In this general discussion, I will summarise the main findings of this study and highlight the questions that also came from the data obtained and finally I propose some further studies to try to answer to these questions.

Variability in growth and hormone content in 17 Arabidopsis accessions

To analyze the variability in growth between the 17 accessions used in this study, plants were grown for 25 days and thirteen leaf growth related parameters were measured. We observed a wide range of natural variation in leaf-growth at plant, organ, and cellular level in these 17 wild-type accessions (Chapter 5). As expected, rosette size positively correlated with the number of leaves, fresh weight, dry weight and leaf area. A positive correlation also was found between leaf area and pavement cell area, whereas no correlation was detected for leaf cell number. The finding that pavement cell area correlates with leaf area but also fresh weight is surprising as detailed characterization of genes that enhance organ size mainly were shown to affect cell number (Mizukami and Fischer, 2000; Autran et al., 2002; Rodriguez et al., 2010). Only few genes including *EXPANSIN 10* (Cho and Cosgrove, 2000) or *SAUR 19* (Spartz et al., 2012) are known to enhance leaf size by promoting cell expansion. However one should not forget that most research on the effect of transgenes on organ

size are performed using the Col-0 accession and that possibly regulation of cell number in this accession is more prominent than in the other accessions used in this study.

We also observed variability in hormone content in the 17 accessions. Several hormones showed a positive correlation with each other highlighting the complex interplay existing between hormones as previously described (Nemhauser et al., 2006). Some links between hormone content and growth related phenotypes were observed, mainly with cellular parameters. However no correlation was found between hormones and whole plant level growth parameters. A possible explanation for no correlation with plant level parameters is that total leaf area, dry weight and fresh weight are complex traits which result from the effects of multiple hormones, in contrast to cellular parameters which might be subject to the influence of one or only a few hormones.

The phenotypic analysis of leaf size related parameters of all wild-type plants and transgenic lines was performed using the in-house phenotyping platform called MIRGIS which is an automated image analysis system for in soil grown plants. After growing the plants for 25 days, leaf series were made to measure each individual leaf area. Although the imaging analysis of the rosette has several advantages such as providing a non-destructive set of data, detailed time-resolution, and saving time compared to the tedious and time-consuming job of making leaf series, the reliability of the data can be influenced by different factors such as leaf shape and overlap of the leaves during the growth. We performed a comparative analysis of the phenotypes of various genotypes based on image analysis of the rosette (non-destructive) and leaf series data (destructive) (Chapter 8). Since we generated various genotypes in different genetic backgrounds, this genetic diversity creates morphological diversity because the different accessions have different periods of vegetative growth, different leaf size and shape, and different number of leaves. This morphological diversity causes that image analysis of the rosette can sometimes be incorrectly estimated or that relevant information is missed. Therefore, in order to obtain accurate measurements of the rosette area, the variation of the leaf shape, the area and the developmental stages of the plant need to be taken into consideration.

In summary, through different levels of phenotyping analysis, we observed a wide range of variability on growth as well as hormone content in 17 Arabidopsis accessions. The analysis of growth parameters and hormone levels in a larger number of accessions (>100) could provide more insight into how leaf growth is regulated in a broader set of natural variants.

The effects of three transgenes in natural variants: epistatic interactions

To introduce a gene from one species into another, introgression method, corresponding to the transfer of a gene by repeated backcrossing of an interspecific hybrid with one of its parent species, for example, has been commonly used for breeding of crops. Using introgression method provides the advantage that the gene of interest will always be surrounded by the same genomic region. However, since introgression is a long process with many hybrid generations, this method could not be used in this project. Therefore transformation method was used for all the accessions. Then there might be a position effect of transgene in different independent transgenic lines within an accession. However, by averaging the effect of different independent transgenics per accession, expressivity of the transgene per accession was possible to estimate.

We conducted a meta-analysis with the expressivity values of the three transgenes introduced in the 17 accessions to compare their effects in different genetic backgrounds (Chapter 7). We found

that the effects of *GA20ox1*^{OE} in the 17 Arabidopsis accessions were more diverse compared to the effects of *amiPPD* and *DN-DA1*^{OE} transgenic lines. Probably, it is because *GA20ox1*^{OE} might affect more broad developmental processes compared to the other transgenes as observed at cellular level. In general, comparative analysis of the effect of each transgene in the different genetic backgrounds revealed that the expressivity is depending on both transgene and background. However, the phenotypic differences are likely not due to sequence diversity between the transgenes and the endogenous genes in the different accessions. Furthermore, the correlation analysis between wild type phenotypes and the expressivity of the transgenes showed different results depending on the transgenes. While the expressivity of *GA20ox1*^{OE} positively correlated with growth related parameters of wild type (biomass, rosette area and/or total leaf number), these correlations were negative for *amiPPD* and *DN-DA1*^{OE}*.* These findings demonstrate that there is no common rule explaining how a wild-type plant will be affected by a transgene, but the outcome could be dependent on epistatic interactions between the transgene and the genetic background and/or differences in the signalling pathway in which the gene of interest is involved. Through correlation analysis between transcript levels in wild type and expressivity of *GA20ox1*^{OE} and *amiPPD* transgenes, we also found that natural variation of basal expression of some genes in wild type might influence the effect of the transgene. For example, the basal expression level of genes involved in ethylene signalling pathway were shown to negatively correlate with rosette expressivity of *GA20ox1*^{OE}. Since ethylene is known to suppress accumulation of bio-active GA forms (Achard et al., 2007; Sun, 2011), an accession which has low expression of gene involved in ethylene biosynthesis or signalling pathway could have high expressivity due to less antagonistic action of ethylene towards GA.

Also it might be that the transgenes have different epistatic interactions, in function of the accession, with other growth regulatory genes which could explain the variation in expressivity observed. In Col-0 background, an epistasis study using binary combinations of growth-promoting genes showed that 39% of the combinations had an additional increase in leaf size mainly resulting from a positive epistasis on growth (Vanhaeren et al., 2014). In function of the genes combined, different size phenotypes were observed. Therefore, it could be that in different accessions depending on the expression of certain growth regulators, the effect of the transgene might be different. In *Drosophila,* by studying how different genetic backgrounds influence the effect of the *scalloped*^{E3} allele on wing shape, a number of genes with a role in wing development as well as several genes with no function in wing development were identified as modifiers affecting the mutant phenotype (Chandler et al., 2014). In our study, other factors that are not involved in GA, PPD or DA1 signalling could act as modifiers that influence the effect of the transgenes in different accessions.

The effect of the transgenes on leaf growth depends on the genetic background

- **Overexpression of *GA20ox1***

From the overexpression of *GA20ox1* in 17 accessions (Chapter 5) several observations could be made. First, most accessions visibly responded by changing their growth especially with an altered leaf size and shape. However, across the accessions, the response did not always correspond to a

positive effect on growth. Eleven accessions showed larger rosettes whereas others had smaller rosette size compared to wild type. Similarly, depending on the genetic background the effect of *GA20ox1*^{OE} in leaf growth at cellular level was different. Since GA controls organ growth by regulating cell division and/or expansion (Achard et al., 2008; Achard et al., 2009; Ubeda-Tomás et al., 2009; Gonzalez et al., 2010; Claeys et al., 2012; Nelissen et al., 2012), it is possible that in different accessions the effect of GA on cell division and/or expansion can be differentially regulated. If in one accession some genes inhibiting cell division are more present, it might be that the effect of increased GA levels is different from other accessions. However, in all transgenic lines, GA levels showed the same direction of accumulation suggesting that GA biosynthesis/metabolism pathway is commonly changed across the accessions. Remarkably, transcript levels of *GA20ox1* do not correlate with the levels of bioactive GA. Furthermore, the levels of bio-active GA forms in the different transgenic lines were remarkably constant for all transgenics in each accession suggesting the existence of an accession-specific plateau for maximal accumulation of these GAs. GA levels were therefore not correlated with the phenotypes suggesting that high accumulation of GA is not always responsible for a positive growth regulation. We also found that biomass of wild type was positively correlated with the observed growth-promoting effects of *GA20ox1*^{OE} on rosette size (rosette expressivity). In other words accessions that are already larger compared to other accessions are more susceptible to show a positive effect of *GA20ox1*^{OE} than smaller accessions. However, no correlation between rosette expressivity and GA levels of wild type was found, whereas selective leaf expressivity was positively correlated with GA_4 levels.

Then, how to explain the variability in phenotypes caused by overexpression of *GA20ox1*? One possible reason that could explain the lack of correlation between *GA20ox1* transcripts or GA_4 levels in the transgenics and the observed phenotypes would be differential control of the GA-related pathway at post-transcriptional level in the different accessions. Translatome analysis after treatment with bioactive GA revealed, for example, that genes involved in feedback regulation of GA showed differential mRNA translation upon treatment (Ribeiro et al., 2012). This differential translation also could occur at varying levels in the accessions. To verify this hypothesis, some accessions could be chosen such as Ler-0 showing high accumulation of GA and highly expressed *GA20ox1* but without positive effects on leaf size and ICE75, as contrast, showing high expressivity with high accumulation of GA and high expression of *GA20ox1.* These accessions could be treated with GA and their translatome compared.

DELLAs, negative regulators of GA signalling, act as transcription regulators (Sun, 2011) and interact with numerous proteins to regulate them post-translationally (Claeys et al., 2014). It is therefore possible that the basal level of DELLA proteins differs in the different accessions explaining the variability in the phenotypic response upon overexpression of *GA20ox1*. It would be interesting to analyse these basal levels in the different wild type accessions and also their degradation in the *GA20ox1-OE* transgenic lines to verify this potential explanation. Additionally, variable activity of the GA-GID-DELLA regulatory module in the different accessions could affect the response of the accessions to an increase level of GA due to the genetic perturbation imposed. The amount of the GA-receptors (GID) and their affinity and efficiency of interaction to form the GA-GID-DELLA complex might be different in the different wild type accessions and therefore cause the differential response observed between the accessions. The sequences of the GID and DELLA could be compared to check if

there are differences which could make the interaction affinity different. Also the level of GID can be compared between the accessions in both wild type and transgenic plants.

Since the levels of bio-active GAs in the transgenic plants showed accession-dependent maximum of accumulation, treatment with variable amounts of gibberellin of the wild type accessions would be also interesting to do to verify if there are different threshold levels in the morphological response of these plants to GA. Few accessions could be selected for this analysis. For example, transgenic plants of two accessions, Cvi-0 and C24, both had low accumulation of GA_4 but showed a contrasting morphological response, high expressivity and low expressivity, respectively. The transgenic plants of two other accessions, ICE75 and Ler-0, have high concentration of GA_4 however, ICE75 showed high expressivity while Ler-0 have low expressivity.

We observed not only common but also accession-specific molecular responses in the transgenic plants. For example, genes involved in photosynthesis showed a strong down-regulation only in some accessions such as WalhaesB4 and Ey15-2. This result suggests that there might be different activity of transcription regulators in these accessions or sequence differences at promoter regions of those accession-specific regulated genes. It has been shown that mutations which influence the function of promoter regions contribute to phenotypic diversity within and between species (reviewed in Wray, 2007; Wittkopp and Kalay, 2012). In yeast, the divergence in the binding sites of specific transcription factors, pseudohyphal regulators Ste12 and Tec1, leads to differences in transcription regulation across the species (Borneman et al., 2007). Therefore, to verify the hypothesis in our study, those accession-specific regulated genes could be selected based on RNA seq data. The sequence of the promoter regions of these genes could be compared between the accessions to check if there is a mutation which could lead to reduced affinity to transcription regulators including DELLAs.

In this study, we found that high accumulation of gibberellin in various natural Arabidopsis accessions did not always positively affect leaf growth and the effect also varied greatly in function of the genetic background. Using these detailed comparative analysis with specific natural variants, we would obtain a better knowledge on not only natural variation but also novel genetic interactions involved in GA signalling.

- **Down-regulation of *PPD2***

In the case of the *amiPPD* transgenic lines in different Arabidopsis accessions (Chapter 6), most of the *amiPPD* transgenic lines were positively affected in the size of their older leaves. But, not all of the transgenic lines showed larger rosette areas. The phenotypic differences seem not to be due to sequence diversity between the transgenes and the endogenous gene in the different accessions. It has been previously shown that down-regulation of *PPD* leads to the formation of larger leaves with a dome-shape in Ler-0 (White, 2006) and Col-0 background (Gonzalez et al., 2015). Interestingly, we found that transgenic plants of two accessions, C24 and Sha, showed increased leaf area without having an obvious dome-shape phenotype while ICE163 displayed a clear dome-shape phenotype with no alteration in leaf size. These results suggest that effects of leaf area and shape are regulated in an independent manner. We observed that an increase in pavement cell number is the main contributor for the enlarged leaf area after down-regulation of *PPD2*. However, no increased pavement cell number was observed in the transgenic plants of ICE163 while Sha had increased both pavement cell number and area in the transgenic plants. These two accessions showed a contrasting phenotype upon down-regulation of *PPD2* at leaf and cellular levels. It would be interesting to compare in more details

the difference at molecular level between these two accessions. Since in this study RNA sequencing analysis was conducted with one repeat, we faced difficulties to isolate significantly differentially expressed genes in a specific accession triggered by the transgenes. Therefore additional transcriptome analysis can be performed in ICE163, Sha and Col-0 as a reference. If PPD2 influences leaf size and shape by regulating different targets, it can be speculated that in ICE163 down-regulation of *PPD2* only affects the targets for leaf shape not for size and in the case of Sha other targets would show a distinct differential expression.

We found no correlation between the degree of down-regulation of *PPD2* and changes of leaf area in the transgenic lines. Since final leaf size is a complex process regulated by a large amount of genes it is possible that epistatic interactions between the genetic background and the genetic perturbation lead to differences in phenotypes. For example, many genes involved in the regulation of cell proliferation and/or cell expansion might have variable activities in the different genetic backgrounds. Therefore buffering or increasing the effect of the increase of meristemoid division by the down regulation of *PPD2* can be triggered. We also observed accession-specific differential expression between the accessions of 30 direct target genes of PPD2 identified in Col-0. A possible reason to explain this accession-specific differential expression would be sequence variants in PPD2 protein as well as promoter regions of the target genes of PPD2. Since we observed PPD2 protein sequence variation in some accessions, this variation might cause differences for the affinity to other co-transcription regulator proteins which PPD2 interacts to function properly. The changes in affinity to other interacting proteins could result in differential regulation of target gene expression. Also it might be that variations in sequence at the promoter region of target genes change the affinity for PPD2 binding which results in differential regulation of the target genes in different accessions. To verify this hypothesis, the sequence of the promoter regions of PPD2 targets in different accessions can be compared to find potential polymorphisms. Since PPD2 requires at least two interacting proteins, KIX8 and KIX9, to regulate target gene expression (Gonzalez et al., 2015), protein sequence as well as the basal levels of these proteins could be checked in the different accessions. Different basal levels of these proteins could explain why we did not observed a clear correlation between down-regulation of *PPD2* and expressivity. For example, in transgenic plants of two accessions having similar levels of PPD2 but different levels of KIX8 or KIX9 which is still lower than the PPD2, the accession which has the lowest level of KIX protein could have the highest expressivity of *amiPPD*.

We observed that some of the direct target genes of PPD2 were down-regulated in the transgenic plants. Since PPD2 contains a tify domain which could act as transcriptional activator (Shikata et al., 2003) and also the KIX protein domain is present in several transcriptional coactivators (Radhakrishnan et al., 1997), we speculated that PPD2 might also act as a transcription activator. To test this one can select few genes that are down-regulated. By using a transient expression assay in protoplasts with promoter-luciferase reporter constructs, this hypothesis can be verified in different accessions for example, in ICE163 and Oy-0 showing down-regulation of many target genes of PPD2 compared to other accessions. In addition, *PPD2* overexpression lines can be generated to check if the down-regulated of target genes in *amiPPD* lines would be up-regulated.

Another possibility to explain the accession specific down- or up-regulation of direct target genes of PPD2 could be that there might be other transcription regulators that influence the expression of the same target genes with different activities in the different accessions. For example, it is possible that in some accessions transcription activators which govern similar targets as PPD2 can be more

active than in other accessions resulting in a higher up-regulation of target genes when PPD2 is absent. A possible way to identify genes that might be responsible for differential expression of target genes which could lead to phenotype variability upon down-regulation of *PPD2* in the different accessions would be conducting a mutation screening in the transgenic lines treated with ethyl methane sulfonate (EMS) that induces random mutations in the genome (Qu and Qin, 2014). By performing a forward genetic screen on specific transgenic lines showing interesting phenotypes, such as ICE163 and Sha, mutations resulting in a positive effect on leaf growth could be identified. From a mutant screening in ICE163, for example, we could look for plants having no dome-shape or increased leaf area phenotype. In the case of Sha, mutants showing dome-shape phenotype or no alteration in leaf size could be selected. To confirm the effect of identified modifier genes in different accessions, CRISPR/Cas 9 system can be used (Fauser et al., 2014) since T-DNA insertion lines are only available in Col-0.

- **Overexpression of a dominant negative form of *DA1***

For the overexpression of the dominant negative form of *DA1* in 17 accessions (Chapter 7), most of the transgenic lines have rounder leaves than their corresponding wild type, similar to what has been previously described in a Col-0 background (Li et al., 2008). We observed an increased leaf area in most of the transgenic lines mainly in older leaves and a decrease in size of the younger leaves at the timepoint analysed. While selective leaf expressivity of this transgene was relatively high, rosette expressivity values are variable in function of the background. At cellular level, increased leaf 3 area in the transgenic lines was mainly due to an increased in pavement cell number and the degree of the effect was depending on the genetic background. However, sequence diversity of DA1 seems to not be a reason of the phenotype differences between the accessions.

DA1 is involved in protein degradation. By modulating the stability of its targets, for example, TCP15 and TCP14 (Peng et al., 2015) and UBP15/SOD2 (Du et al., 2014), plant organ size is regulated. A possible reason that could explain the various effects of *DN-DA1*^{OE} in the different accessions might be that there are different targets and activity of DA1 in different accessions. Therefore, proteomics could be conducted in, for example, Blh-1 and Ey15-2 having high and low expressivity of *DN-DA1*^{OE} respectively, to compare the downstream targets of DA1. When a dominant negative form of DA1 is expressed, due to a competition between the endogenous protein and the mutant form, target proteins of DA1 are expected not to be properly degraded. Western blot could be conducted to check the level of these proteins in these two accessions. Furthermore, protein sequences of UBP15/*SOD2,* TCP15 and TCP14, could be determined as a possible reason for the different effect of DA1 in Blh-1 and Ey15-2. Basal level and fold change accumulation of these proteins in transgenic plants could be analyzed to see if stability of the proteins affects penetrance. Especially, since TCP15 and TCP14 are involved in the regulation of pavement cell area by influencing endoreduplication, the level of these proteins could also be checked in ICE163 and Sha that showed alteration of pavement cell area in the transgenic plants. Also, in Col-0, DA1 interacts with DA2, an ubiquitin ligase, (Xia et al., 2013) to regulate organ size. Since overexpression of *DA2* leads to formation of smaller leaves compared to wild type, basal level of DA2 can be analyzed in Blh-1 and Ey15-2. It might be that Ey15-2 has higher accumulation of DA2. To check if sequence variation of DA2 might cause differences of affinity with DA1, the protein sequences of DA2 should also be analyzed in Blh-1 and Ey15-2.

In summary, we observed various morphological response of *DN-DA1*^{OE} in the different accessions. Regarding the function of DA1, further analysis related to targets in different backgrounds could bring more insight into natural variation response of this genetic perturbation.

Conclusion

In conclusion, we found that phenotypic outcome by genetic perturbation varies substantially depending on the combination between transgene and genetic background probably as a result of epistatic interactions. The genetic network might be flexible depending on the interaction with a number of different factors influencing the phenotype outcomes in different backgrounds (Figure 9.1) (Cooper et al., 2013). By using the generated transgenic plants for the experiments described in this discussion, knowledge on these epistatic interactions will not only provide insights in the molecular basis of genetic variation but also will generate insightful knowledge on genetic networks and phenotypic robustness. Moreover, similar to its use for personalized medicine in human health, such understanding will allow for targeted crop improvement to predict penetrance of specific mutations in different genetic background.

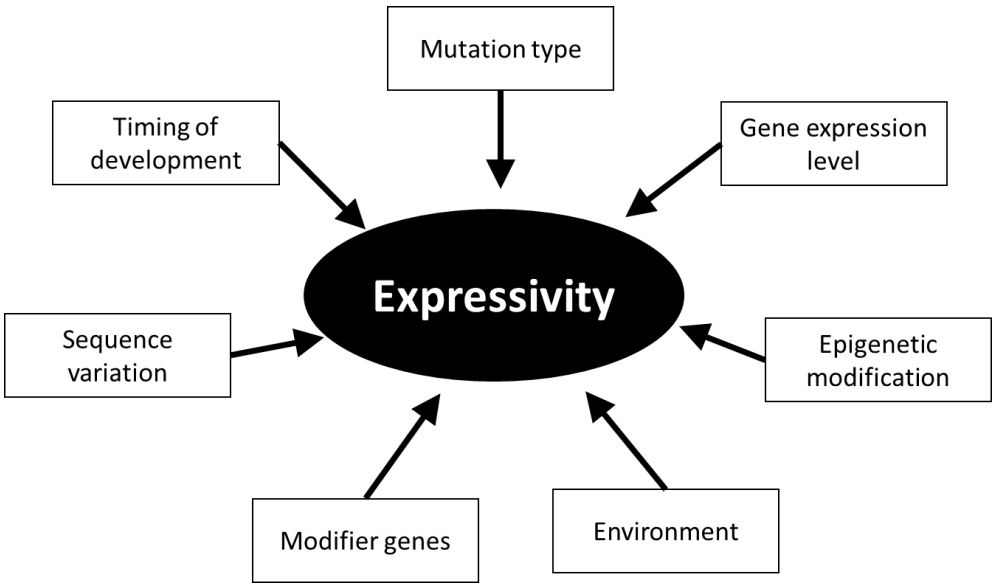


Figure 9.1 Several different factors influencing expressivity of a specific genetic perturbation (modified from Cooper et al., 2013).

References

Achard, P., Renou, J.-P., Berthomé, R., Harberd, N.P., and Genschik, P. (2008). Plant DELLAs restrain growth and promote survival of adversity by reducing the levels of reactive oxygen species. Curr. Biol. **18:** 656-660.

Achard, P., Baghour, M., Chapple, A., Hedden, P., Van Der Straeten, D., Genschik, P., Moritz, T., and Harberd, N.P. (2007). The plant stress hormone ethylene controls floral transition via DELLA-

dependent regulation of floral meristem-identity genes. Proc Natl Acad Sci U S A **104:** 6484-6489.

Achard, P., Gusti, A., Cheminant, S., Alioua, M., Dhondt, S., Coppens, F., Beemster, G.T.S., and Genschik, P. (2009). Gibberellin signaling controls cell proliferation rate in *Arabidopsis*. Curr. Biol. **19:** 1188-1193.

Autran, D., Jonak, C., Belcram, K., Beemster, G.T.S., Kronenberger, J., Grandjean, O., Inzé, D., and Traas, J. (2002). Cell numbers and leaf development in *Arabidopsis*: a functional analysis of the *STRUWWELPETER* gene. EMBO J. **21:** 6036-6049.

Borneman, A.R., Gianoulis, T.A., Zhang, Z.D., Yu, H., Rozowsky, J., Seringhaus, M.R., Wang, L.Y., Gerstein, M., and Snyder, M. (2007). Divergence of transcription factor binding sites across related yeast species. Science **317:** 815-819.

Chandler, C.H., Chari, S., and Dworkin, I. (2013). Does your gene need a background check? How genetic background impacts the analysis of mutations, genes, and evolution. Trends Genet **29:** 358-366.

Chandler, C.H., Chari, S., Tack, D., and Dworkin, I. (2014). Causes and consequences of genetic background effects illuminated by integrative genomic analysis. Genetics **196:** 1321-1336.

Cho, H.-T., and Cosgrove, D.J. (2000). Altered expression of expansin modulates leaf growth and pedicel abscission in Arabidopsis thaliana. Proceedings of the National Academy of Sciences **97:** 9783-9788.

Claeys, H., De Bodt, S., and Inze, D. (2014). Gibberellins and DELLAs: central nodes in growth regulatory networks. Trends Plant Sci **19:** 231-239.

Claeys, H., Skirycz, A., Maleux, K., and Inzé, D. (2012). DELLA signaling mediates stress-induced cell differentiation in Arabidopsis leaves through modulation of anaphase-promoting complex/cyclosome activity. Plant Physiol. **159:** 739-747.

Cooper, D.N., Krawczak, M., Polychronakos, C., Tyler-Smith, C., and Kehrer-Sawatzki, H. (2013). Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease. Human genetics **132:** 1077-1130.

Du, L., Li, N., Chen, L., Xu, Y., Li, Y., Zhang, Y., Li, C., and Li, Y. (2014). The ubiquitin receptor DA1 regulates seed and organ size by modulating the stability of the ubiquitin-specific protease UBP15/SOD2 in *Arabidopsis*. Plant Cell **26:** 665-677.

Fauser, F., Schiml, S., and Puchta, H. (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. The Plant Journal **79:** 348-359.

Gonzalez, N., et al. (2010). Increased leaf size: different means to an end. Plant Physiol. **153:** 1261-1279.

Gonzalez, N., et al. (2015). A Repressor Protein Complex Regulates Leaf Growth in Arabidopsis. Plant Cell **27:** 2273-2287.

Huang, S., Raman, A.S., Ream, J.E., Fujiwara, H., Cerny, R.E., and Brown, S.M. (1998). Overexpression of 20-oxidase confers a gibberellin-overproduction phenotype in Arabidopsis. Plant Physiol **118:** 773-781.

Li, Y., Zheng, L., Corke, F., Smith, C., and Bevan, M.W. (2008). Control of final seed and organ size by the *DA1* gene family in *Arabidopsis thaliana*. Genes Dev. **22:** 1331-1336.

Mizukami, Y., and Fischer, R.L. (2000). Plant organ size control: *AINTEGUMENTA* regulates growth and cell numbers during organogenesis. Proc. Natl. Acad. Sci. USA **97:** 942-947.

Nelissen, H., Rymen, B., Jikumaru, Y., Demuynck, K., Van Lijsebettens, M., Kamiya, Y., Inzé, D., and Beemster, G.T.S. (2012). A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. Curr. Biol. **22:** 1183-1187.

Nemhauser, J.L., Hong, F., and Chory, J. (2006). Different plant hormones regulate similar processes through largely nonoverlapping transcriptional responses. Cell **126:** 467-475.

Peng, Y., Chen, L., Lu, Y., Wu, Y., Dumenil, J., Zhu, Z., Bevan, M.W., and Li, Y. (2015). The ubiquitin receptors DA1, DAR1, and DAR2 redundantly regulate endoreduplication by modulating the stability of TCP14/15 in Arabidopsis. Plant Cell **27:** 649-662.

Qu, L.-J., and Qin, G. (2014). Generation and identification of Arabidopsis EMS mutants. In Arabidopsis Protocols (Springer), pp. 225-239.

Radhakrishnan, I., Pérez-Alvarado, G.C., Parker, D., Dyson, H.J., Montminy, M.R., and Wright, P.E. (1997). Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: a model for activator:coactivator interactions. Cell **91:** 741-752.

Ribeiro, D.M., Araújo, W.L., Fernie, A.R., Schippers, J.H., and Mueller-Roeber, B. (2012). Translatome and metabolome effects triggered by gibberellins during rosette growth in *Arabidopsis*. J. Exp. Bot. **63:** 2769-2786.

Rodriguez, R.E., Mecchia, M.A., Debernardi, J.M., Schommer, C., Weigel, D., and Palatnik, J.F. (2010). Control of cell proliferation in *Arabidopsis thaliana* by microRNA miR396. Development **137:** 103-112.

Shikata, M., Takemura, M., Yokota, A., and Kohchi, T. (2003). Arabidopsis ZIM, a plant-specific GATA factor, can function as a transcriptional activator. Biosci. Biotechnol. Biochem. **67:** 2495-2497.

Spartz, A.K., Lee, S.H., Wenger, J.P., Gonzalez, N., Itoh, H., Inzé, D., Peer, W.A., Murphy, A.S., Overvoorde, P.J., and Gray, W.M. (2012). The *SAUR19* subfamily of *SMALL AUXIN UP RNA* genes promote cell expansion. Plant J. **70:** 978-990.

Sun, T.-p. (2011). The molecular mechanism and evolution of the GA-GID1-DELLA signaling module in plants. Curr. Biol. **21:** R338-R345.

Ubeda-Tomás, S., Federici, F., Casimiro, I., Beemster, G.T.S., Bhalerao, R., Swarup, R., Doerner, P., Haseloff, J., and Bennett, M.J. (2009). Gibberellin signaling in the endodermis controls *Arabidopsis* root meristem size. Curr. Biol. **19:** 1194-1199.

Vanhaeren, H., Gonzalez, N., Coppens, F., De Milde, L., Van Daele, T., Vermeersch, M., Eloy, N.B., Storme, V., and Inze, D. (2014). Combining growth-promoting genes leads to positive epistasis in Arabidopsis thaliana. eLife **3:** e02252.

White, D.W.R. (2006). *PEAPOD* regulates lamina size and curvature in *Arabidopsis*. Proc. Natl. Acad. Sci. USA **103:** 13238-13243.

Wittkopp, P.J., and Kalay, G. (2012). Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. Nat. Rev. Genet. **13:** 59-69.

Wray, G.A. (2007). The evolutionary significance of cis-regulatory mutations. Nat Rev Genet **8:** 206-216.

Xia, T., Li, N., Dumenil, J., Li, J., Kamenski, A., Bevan, M.W., Gao, F., and Li, Y. (2013). The ubiquitin receptor DA1 interacts with the E3 ubiquitin ligase DA2 to regulate seed and organ size in *Arabidopsis*. Plant Cell **25:** 3347-3359.

Summary

The relationship between a specific genetic perturbation and a phenotype depends not only on the environment, but also on the genetic background in which the perturbation occurs. Variation in the effect of a mutation in different genetic backgrounds, caused by genetic interactions, has been observed in several organisms including bacteria, yeast, worm, fruit fly, and mouse. However, the knowledge about how genetic background influences the phenotypic output of a mutation in plants is poorly understood. The genetic interactions for quantitative traits such as leaf growth in plants are difficult to identify due to the complexity of this process. Numerous growth-promoting genes have been identified that regulate leaf growth in Arabidopsis. To functionally analyse these genes, genetic perturbations are often carried out in one genetic background for instance the Columbia-0 background (Col-0) in Arabidopsis. However, a stable effect of a transgene across different genetic backgrounds is necessary to have robust phenotype with high expressivity. Therefore, understanding of mechanisms that allow for predict and control the expressivity is required.

In this project, we examined natural variation of molecular and morphological response to three transgenes known to positively regulate leaf growth in Col-0 background: *GA20ox1-overexpression* (*GA20ox1*^{OE}), overexpression of an artificial microRNA targeting *PPD (amiPPD)*, and overexpression of a dominant negative form of *DA1 (DN-DA1*^{OE}*)*. These three transgenes were introduced in seventeen Arabidopsis natural accessions from around the world. First, the variability of growth related parameters in the 17 accessions was analyzed at plant, organ, and cellular level. Second, the leaf size of all transgenic lines was measured with leaf series. Third, transcriptome analysis was conducted to examine the cause of the phenotypic changes at molecular level in the transgenics.

In general, in the majority of the transgenic lines, growth promoting effects were observed. However, the expressivity varied depending on the accession and the transgene. For the lines overexpressing *GA20ox1*, younger leaves were positively affected whereas older leaves showed increased leaf area in both *amiPPD* and *DN-DA1* ^{OE} transgenics compared wild type. Cellular analysis on *GA20ox1* ^{OE} transgenics revealed that different cellular changes caused a change in final leaf size in function of the accession analysed. We found different correlations between wild-type size related-parameters and the expressivity of the transgenes. Biomass of wild-type accession was positively correlated with the expressivity of *GA20ox1*^{OE} transgene, but negatively correlated with the expressivity of *amiPPD* transgene. Differentially expressed (DE) genes in at least one accession were identified by transcriptome analysis for *GA20ox1* ^{OE} and *amiPPD* transgenic lines. From the correlation analysis, we identified DE genes correlated with the expressivity of the transgenic lines.

In this project, through comparative analysis using multiple transgenes and accessions, common and accession-specific responses to genetic perturbation were observed in 17 Arabidopsis accessions.

Acknowledgement 마치는 글

이 논문 중에서 가장 쓰기 힘든 파트 (글도 쓰기 전에 눈물부터 나오는 파트!) 라고 해도 과언은 아니다. 몇 일을 걸렸지만 아직도 한 단락도 채 쓰지 못했다. 쓰기 싫은게 아니라(!), 내가 보낸 그 시간들을 정리하는 기분이라, 어떤 말로 써도 부족한 느낌이 들었기 때문이다. 4 년 8 개월이라는 긴 시간을, 이 논문은 1% 도 채 보여주지 못하는거 같다. 그 긴 시간 벨기에에서 공부를 하는 동안 많은 것을 느끼고, 생각하고, 혼자 있는 시간을 즐기고 또 괴로워 하면서 나는 부쩍 큰 느낌이다. 올해 32 살이 되었지만 여전히 학생이라는 신분을 만끽 하면서 생활했던 이 시간들을 이제 마치려고 한다. 그 동안 도와준, 함께 해준 사람들에게 이렇게 나마 고마움을 전하고 싶다.

This is the last part of this thesis I write and I would say it's probably the most difficult part to write in this thesis. Because there are so many people I would like to thank and so many thought, feelings, and memories I had during 4 years and 8 months in Belgium.

First of all, I would like to thank Dirk for giving me a chance to do PhD in PSB and supporting me during my PhD. I've learned from you what is "enthusiastic". ☺ And I would like to appreciate co-promoter of my PhD (also you're my friend!), Nathalie. Hey, here we are, the end! Finally. I'm really glad that I've worked with you.

Hey, Yielders (Alexandra (Ola), Hannes C, Hannes V, Marieke, Hilde, Jonas, Jasmien, Ting, Chrystalla, Yiota, Lisa, Alexandra, Mattias, Katrien, Twiggy, Judith, Kirin, Takasaki, Raf, Dorata, Frederik, Jasper, Stijn, Liesbeth V, Liesbeth M, Charlot, Lennart, Stien, Kim, Jolien, Kaoru, Nathalie W, and Nathalie V) thanks a lot for helping me and having a nice discussion as well as lots of beers! You are, because we are!

My friends, Xiao-huan, Alex and Kun (can't imagine my life here without you guys. you guys are like my family in Belgium, love you guys so much!), Rossella (my dearest stronza!), Juliana, Luiz, Igor, Livia, and Nubia (my favorite Brazilians!), Feng, Dalong, Dorien, Tom V and Brecht.

한국 사람 찾기 힘든 이 벨기에 겐트에서, 같이 모여서 한국음식 해먹고 또 한국말 잊지 않게 해준 ^-^, 영주언니, 민정언니, 진영언니, 민자언니, 조박사님, 지은언니, 미영언니, 셀라, 그리고 루벤에서 만난한국분들(보라, 지하, 순길 선배). 마지막으로 멀리 독일에서 논문 심사하시러 두번이나 흔쾌히 와주신 은영 박사님, 무지 감사드립니다.

사랑하는 우리 가족, 엄마, 아빠, 오빠. 내가 한국에 없던 지난 근 7 년 동안, 항상 나를 응원해주고 걱정 해줘서 고마워. 생신이나 어버이날, 결혼기념일엔 매번 오빠만 챙겨줘야 해서 많이 미안했어 (오빠, 고마워!). 이제 나 한국가면 여행 많이 다니고, 내가 요리 많이 해줄께. 모두 건강하고 더욱 더 행복해 지자. 사랑해!

이 녀석(이라 쓰고 논문이라 읽음)을 쓰기 위해, 정말로 x 100 힘들었지만, 그 힘든 와중에 무엇이 나를 즐겁게 하고 행복하게 해주는지 확실히 알수 있었다. 꼭 이런식으로 알고 싶지 않아도 되지만, 여튼 깨닫게 해주어서 감사하다~ 올해를 맞으면서 나는 아주 설레였다. 아무런 계획이 없다는 말은 다른 말로는, 어떤 것도 할수 있다는 무한한 가능성을 보여주고 있으니까.

아, 결국엔 이렇게 짧게 쓸거 였으면서 뭐가 그렇게 쓰기 힘이 들었는지 (생각이 많은게 문제였어..). 결국엔 감사의 글은 나의 넋두리로 끝이 나는 것인가(!). 논문 전체는 scientific 하게 쓰여져 있지만, 끝맺음은 온전히 나 답게 쓰고 싶은거 였나 보다 (나는 진지한것이 어울리지 않는것!). 항상 마지막은 이런 기분이다. '시원섭섭'이라는 말은 누가 만들었는지 모르지만 기가 막히게 이 상황을 잘 표현해 주고 있다. 서운하지만, 기분좋게 마무리 할수 있어서 행복하다. 인생은 시트콤이라는 내 친구의 말을 다시 되새기면서, 늦게 시작해 늦게 끝난 나의 박사 졸업이, 힘들었지만, 너무나 소중한 경험과 시간들 이었다고 마지막엔 웃으며 벨기에를 떠난다.

안녕, 벨기에 그 동안 즐거웠어!

Good bye, Belgium! It was fun!

Curriculum Vitae

Youn-Jeong NAM

Date of birth: February 6, 1984
Place of birth: Cheongsong, Kyungsangpookdo, South Korea
Nationality: South Korean

☎ +32 (0) 48 3509 186
✉ yonam@psb.vib-ugent.be
younjeong.nam@gmail.com

Education

2011/08 – 2016/04 **Ph.D Degree** in Dept. of Plant systems biology from Gent Univ. Belgium.
-thesis title; Natural variation of morphological and molecular responses of genetic perturbations in Arabidopsis.

2007/09 – 2009/08 **Master Degree** in Dept. of Molecular & Cellular biology from Korea Univ. Korea.
-thesis title; Characterization of tobacco IspH specifically interacting with cucumber mosaic virus-encoded 2b protein.

2002/03 – 2007/02 **Graduated** in Dept. of Plant Bioscience & Life Science and Biotechnology from Kyung-pook National Univ. Korea.
- thesis title; The effect of cold-treatment on germination rate in spinach seed

Professional Experience

2009/09 – 2011/04 **Visiting researcher**, RIKEN Plant Science Center, Yokohama, Japan.
Environmental Response Adaptation Research, Regulatory Network Research Unit
- selected as a recipient for the global internship program by the National Research Foundation (NRF) of Korea
1. Relationship between 14-3-3 and WRKY TF in *Arabidopsis*
2. Roles of cytokinin in low-potassium signaling in *Arabidopsis*

2007/03 – 2007/08 **Assistant researcher,** Plant Signaling Network Research Center, Seoul, Korea.
Study about the function of CaBtf3 using Post-transcriptional Gene silencing in *capsicum annum*

2007/01 – 2007/02 **Assistant researcher,** Rural Development Administration, Suwon, Korea.
Making cDNA library of *Capsicum annum*

2006/09 – 2006/12 **Assistant researcher,** School of Agricultural Biotechnology and Center for Agricultural Biomaterials, Seoul National Univ. Korea
Overexpression of Arabidopsis Tocopherol Cyclase in Lettuce

2006/06 – 2006/08 **Assistant researcher,** Rural Development Administration, Suwon, Korea.
Making biological marker in *Nicotiana tabaccum*

Technical Skills

Biochemistry	Proteins (Extraction method, SDS-PAGE, Western blotting, Transient Expression)
Molecular Biology	DNA work (Cloning, PCR, genomicDNA prep, Primer design, Southern blotting); RNA work (RNA prep, RT-PCR, Realtime-PCR, Northern blotting, Gene Silencing using PTGS assay in hotpepper and tobacco); Protein assay (Yeast-two hybrid assay, BiFC assay, Co-localization assay); Infiltration assay in tobacco; Selection of Salk line
Cell Biology	Observation plant and making image with Confocal Laser Microscope and Disection Microscope
Plant Phathology	Pathogen (Virus, Bacteria, and Fungus) Treatment and Harvest
Plant Physiology	Plant growth observation; Root Growth assay; Potassium growth assay; ROS Observation, Chlorophyll measurement
Bioinformatics	Statistical analysis program (Graph Pad Prism, SPSS statistics); Image J); Gene Vestigator; Vector NTI; DNA Star; Photoshop; Microsoft Office
Field Work	Planting

Poster submissions

2015/07	***France***	Natural variation in the effect of GA20ox1- overexpression in Arabidopsis - The 26th International Conference On Arabidopsis Research
2015/03	***Belgium***	Natural variation in the effect of GA20ox1- overexpression in Arabidopsis - Plant organ growth symposium
2012/10	***Korea***	Natural variation in leaf growth rate in Arabidopsis - 10th International Congress on Plant molecular biology
2010/06	***Japan***	AHK Cytokinin Receptor Histidine Kinases Regulates Low Potassium Signaling in *Arabidopsis* - The 21th International Conference On Arabidopsis Research
2009/05	***Korea***	Characterization of tobacco IspH specifically interacting with *cucumber mosaic virus*-encoded 2b protein - The 66th KSBMB Annual Meeting
2007/09	***Korea***	Nuclear proteome analysis of tobacco mosaic virus-induced hot pepper plants: funtional study of 26S proteasome subunit PRN7 - The 4th Solanaceae Genome Workshop

Publication

Youn-Jeong Nam, Lam-Son Phan Tran, Mikiko Kojima, Hitoshi Sakakibara, Rie Nishiyama, Ryoung Shin (2011) Regulatory Roles of Cytokinins and Cytokinin Signaling in Response to Potassium Deficiency in Arabidopsis. PLoS ONE 7(10): e47797

Awards

2009 - 2010 Global Internship Program by the National Research Foundation (NRF) of Korea
2007 - 2008 Graduated Student Scholarship by Brain Korea 21
2007 The Best College Graduation Student by Kyung-pook National Univ.

Language

	Speaking	Writing
Korean	Native	Native
English	Good	Good
Japanese	Basic	Basic